

317471

16/1992

Aikalmazott matematikai lapok

10.

1992/1-2

A MAGYAR TUDOMÁNYOS AKADEμία
MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

16.

KÖTET

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

BENCZÚR ANDRÁS

FŐSZERKESZTŐ-HELYETTESEK

DEMETROVICS JÁNOS, FARKAS MIKLÓS

FELELŐS SZERKESZTŐ

SZÁNTAI TAMÁS

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csizsár Imre, Galántai Aurél, Gécség Ferenc, Gyires Béla, Gyórfy László, Harnos Zsolt, Hatvani László, Heppes Aladár, Kátai Imre, Katona Gyula, Kis Ottó, Klafszky Emil, Kovács Margit, Lovász László, Maros István, Prékopa András, Recski András, Stoyan Gisbert, Tandori Károly, Tusnády Gábor, Varga László

XVI. kötet 1–2. szám

Szerkesztőség és kiadóhivatal: 1117 Budapest, Bogdánfy út 10/B.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus újszerű és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelenítése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Benczúr András, főszerkesztő
1117 Budapest, Bogdánfy út 10/B.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 850 forint. Megrendelések a szerkesztőség címén lehetségesek (átutalásokat az ELTE MNB 232–90142–6207 számlaszámára a 9015567 munkaszám és 694 utalványozási kód megjelöléssel kérjük)

A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungaricae,
2. Acta Physica Hungaricae,
3. Studia Scientiarum Mathematicarum Hungarica.

SÍKBELI, KONZERVATÍV TENZORMEZŐK FORGÁSSZIMMETRIÁINAK VIZSGÁLATA ELEMI MÓDSZEREKKEL*

DOMOKOS GÁBOR

Budapest

A mérnöki gyakorlatban általánosan elterjedt az a feltételezés, amely szerint a *Taylor-sorfejtés* rendjét növelve a vizsgálat eredménye egyre pontosabb lesz, csak kvantitatív információt nyújt.

Vizonylag kevés olyan tétel van azonban, amely a *Taylor-sor* egy adott n rendjéhez a matematikai modellben jelentkező minőségi ugrást rendel hozzá. Talán a legismertebb ilyen jellegű megállapítás, hogy az elsőrendű ($n = 1$) elmélettel nem lehet kritikus terhet, a másodrendű ($n = 2$) tagok figyelembevételével pedig nem lehet posztkritikus viselkedést számolni. Közvetlenül az utóbbi megállapításhoz csatlakozik THOM (1975) tétele. THOM legfeljebb 5 paraméterig osztályozza a potenciális energia függvény kritikus pontjait. Minden esetre megadja a lokális vizsgálathoz (tehát a kezdeti posztkritikus viselkedés meghatározásához) elégséges n értéket.

Ebben a dolgozatban azt mutatjuk meg, hogy a diszkrét forgási szimmetriával rendelkező konzervatív tenzormezők, illetve az ezek által leírt mechanikai jelenségek sok esetben véges sok tagból álló *Taylor-sorokkal* helyettesíthetők anélkül, hogy a szimmetria-csoport rendje megváltozna. Ezzel azt mutatjuk meg többek között, hogy van olyan n , amelyet átlépve a csonkolt sor szimmetriájának rendje megváltozik, amit joggal nevezhetünk minőségi ugrásnak. Az alkalmazások szempontjából érdekes állításokat egy általánosabb tétel korolláriumaként fogalmazzuk meg. Az általános tétel szerint a kétváltozós polinomokat C_n -szimmetrikus magok és körszimmetrikus részek szorzatának lineáris kombinációjaként lehet felírni. A bizonyított tételek alapján meghatározzuk a kétváltozós polinomok véges szimmetria-csoportjai rendjeinek legkisebb felső korlátját. Megmutatjuk, hogy melyek a homogén polinomok illetve a konstans, konzervatív tenzormezők esetén lehetséges forgási szimmetria-csoportok.

1. Bevezetés

A dolgozat célja kettős: egyrészt áttekintést kíván nyújtani a címben jelzett tenzormezők szimmetria-viszonyairól. Ezen azt értjük, hogy összefüggéseket állapítunk meg a tenzormező rendje, az alkalmazott *Taylor-sorfejtés* fokszáma és a diszkrét forgásszimmetria rendje között. A dolgozat ezen első részének a gondolatmenete a következő: A konzervatív tenzormezők vizsgálata visszavezethető az analitikus függvények vizsgálatára, ami viszont a polinomok vizsgálatára redukálható. A dolgozat központi megállapítása a 2.1 lemma illetve a 2.1 tétel, amely a polinomok előállítására egy olyan kettős függvényösszeget ad meg, amely a *Taylor* és *Fourier felbontás* keveréke, és ezáltal lehetővé teszi a két sorfejtés közötti kapcsolat vizsgálatát. Ezt a kettős összeget kettős függványsorra lehet általánosítani az

*A kutatást az OTKA a 683 számú téma keretében támogatta.

analitikus függvények előállításakor. A kettős összeg szoros kapcsolatban van az úgynevezett harmonikus polinomokkal; ezt részletesebben is be fogjuk mutatni.

A dolgozatban bizonyítandó lemmák és tételek szorosan összefüggenek az invariánsok *Hilbert-Mumford-féle* geometriai elméletében (DIEUDONNÉ és CARRELL, 1971) megfogalmazott tételekkel illetve SCHWARZ (1975) megállapításaival, sőt, a dolgozatban megfogalmazott állítások egy része a *Hilbert-Mumford elmélet* általános tételeinek korolláriumaként is tekinthető. Mindazonáltal a dolgozat alapvetően mérnöki szempontból felvetődő problémákra keresi a matematikai választ, és ezen problémákat a matematikusok fel sem vetették, ezért nem is válaszolhattak rájuk. További célja a dolgozatnak, hogy a tételeket elemi eszközökkel bizonyítsa, és ez a törekvés — ha más és más szempontból is — a mérnökök és a matematikusok számára egyaránt érdekes lehet. Nem lesz tehát szükség a *Hilbert-Mumford elméletben* felhasznált csoport reprezentáció-elméletre és más mély matematikai eszközökre.

A dolgozat másik célja a kidolgozott matematikai eredmények alkalmazása ciklikusan szimmetrikus tartószerkezetek mérnöki számítására. A ciklikusan szimmetrikus szerkezetek számításával kapcsolatban számos speciális eredményt érték el és módszert dolgoztak ki. Statikailag határozott rácsostartók rúderőinek meghatározásával foglalkozik HEGEDŰS (1978). A hengeres rácsok statikai és kinematikai jellemzőit vizsgálja TARNAI (1980). Hengeres rácsok dinamikai vizsgálatával foglalkozik TORNÝOS (1985). Hengeres rács példáján illusztrálja eredményeit GÁSPÁR és TORNÝOS (1986).

Ciklikusan szimmetrikus szerkezetek számításánál sokszor az jelenti a gondot, hogy a szerkezet mechanikai jellemzőit leíró függvények (potenciális energia, merevségi operátor, feszültségi tenzormező) csonkolt *Taylor-sora* nem tükrözi a szerkezet, illetve a sorba fejtett függvény diszkrét forgásszimmetriáját, hanem folytonos körszimmetriát, illetve magasabb rendű diszkrét szimmetriát mutat. (Például a négyzet keresztmetszetű rúd másodrendű tehetetlenségi nyomatéka minden súlyponti tengelyre azonos.) A matematikai eredmények alapján azt szeretnénk megvizsgálni, hogy adható-e szükséges és elégséges feltétel a *Taylor-sor* figyelembe veendő tagjainak darabszámára, ha feltételül szabjuk, hogy a csonkolt sor megőrizze az eredeti függvény szimmetria-tulajdonságait, vagyis a csonkolásra mint operációra szimmetria-invariáns legyen. További feladat annak meghatározása, hogy adott rendű szimmetria esetén mi az összefüggés a csonkolt *Taylor-sor* együtthatói között.

2. Tenzorok szimmetria-invarianciája

Ebben a pontban csak síkbeli tenzormezőkkel foglalkozunk.

2.1 Definíció. Konzervatív tenzormezőnek nevezünk egy olyan m -edrendű D_m tenzormezőt, melynek 2^m koordinátája egy $[xy]$ derékszögű koordinátarendszerben értelmezett $z = F(x, y)$ analitikus skalármező 2^m darab, m -edrendű parciális deriváltjával egyezik meg.

A koordináták egy m -dimenziós egységkocka csúcsaiban helyezkednek el, tehát minden koordináta egy m hosszúságú bináris $(0,1)$ elemekből álló sorozattal azonosítható. A parciális deriválás sorrendje felcserélhető, ezért minden olyan csúcspontban, amely k darab 0-ból és $m - k$ darab 1-esből álló sorozattal azonosítható, a

$$\frac{\partial^m F}{\partial x^k \partial y^{m-k}}$$

derivált áll. Tenzormezők koordináta-invariáns tárgyalásával és mechanikai interpretációjával DOMOKOS (1989) foglalkozik behatóbban.

A továbbiakban csak konzervatív tenzormezőkkel foglalkozunk, ezért általában a „konzervatív” jelzőt elhagyjuk. Speciálisan $m = 0$ esetén nulladrendű konzervatív tenzormezőhöz jutunk, mely magával a $z = F(x, y)$ skalármezővel azonos.

2.1 Skalárok

2.2 Definíció. Egy transzformációkból álló halmazt akkor nevezünk csoportnak, ha tartalmazza bármely két hozzá tartozó transzformáció szorzatát, bármely hozzá tartozó transzformáció inverzét és az identitás-transzformációt. A halmazban található különböző transzformációk száma a csoport rendje. (BIRKHOFF és MACLANE, 1965)

2.3 Definíció. Az origó körüli $u_k = \frac{2k\pi}{n}$ szögekkel történő u_k síkbeli forgatásokból álló C_n transzformáció-csoportot n -edrendű ciklikus csoportnak nevezzük (COXETER, 1969). Az origó körüli tetszőleges forgatások csoportját C_0 -al jelöljük.

A fenti definícióval kapcsolatban megjegyezzük, hogy az origó körüli tetszőleges forgatások csoportja alapvetően eltér a definíció első részében tárgyalt véges csoportoktól, melyeknél a természetes számokat használtuk indexként. A tetszőleges forgatások csoportjának indexelésére tehát egy olyan szimbólumot kellett alkalmaznunk, amely nem természetes szám. Kézenfekvőnek tűnhetne a ∞ index, hiszen ez a csoport végtelen sok transzformációt tartalmaz. Ez az index azonban egyrészt félrevezető, hiszen a végtelen nem szám, hanem számosság, viszont a csoport elemeinek számossága nem azonos a természetes számok számosságával, hiszen a csoport folytonos. Másrészt a későbbiekben látni fogjuk, hogy a ∞ index használata formális szempontból nehézkessé tenné ezen csoport együttes kezelését a véges csoportokkal. A 0 index ezzel szemben nem félrevezető, hiszen nyilvánvalóan nem a transzformációk számára ill. számosságára utal, formális szempontból pedig lehetővé fogja tenni a véges és a folytonos csoport együttes leírását.

2.4 Definíció. Az \mathfrak{R} halmazt egy csoportra nézve invariánsnak nevezzük, ha a csoport bármely transzformációjára nézve invariáns. Az \mathfrak{R} halmazt akkor nevezzük egy transzformációra nézve invariánsnak, ha a transzformált halmazt az eredetitől nem lehet megkülönböztetni. Ha létezik olyan n legnagyobb, véges természetes szám, melyre igaz, hogy \mathfrak{R} invariáns C_n -re nézve, akkor ezt a $C(\mathfrak{R}) = n$ módon

fogjuk jelölni. C_n -et \mathfrak{K} szimmetria-csoportjának nevezzük. Ha nem létezik ilyen szám, és feltesszük, hogy \mathfrak{K} zárt, akkor $C(\mathfrak{K}) = 0$ és C_0 -t \mathfrak{K} szimmetria-csoportjának nevezzük.

A fenti definícióval kapcsolatban megjegyezzük, hogy talán matematikai szempontból pontosabbnak tűnne, ha a „nem lehet megkülönböztetni” megjelölés helyett az „azonos” megjelölést használnánk, ekkor azonban definíciókat semmilyen valóságos tárgyra nem lehetne alkalmazni, pusztán annak matematikai modelljére. Mivel azonban a matematikai modellt éppen azért nyerjük a valóságból, hogy a vizsgálat szempontjából azonosan viselkedő elemeket abszolút értelemben is azonosnak tekintünk, ezért definíciónk nem pontatlan, csupán feltételezi ezen modellalkotási lépés megtételét. Talán nem érdektelen megemlíteni, hogy LEIBNIZ (1848) a következő meghatározást alkalmazta: „Két alakzatot akkor tekintünk ekvivalensnek, ha önmagukban szemlélve nem különböztethetők meg egymástól, mivel minden olyan sajátságuk megegyezik, amelynek objektív jelentés tulajdonítható.” Ez a meghatározás mérnöki (fizikai) és matematikai szempontból egyaránt értelmezhető, hiszen csak az „objektivitás” fogalmát kell meghatároznunk. Mérnöki szempontból az „objektív = mérhető” meghatározás kézenfekvő. Matematikai szempontból WEYL (1946) ad igen egyszerű meghatározást: „Ha a geometria egy axióma-rendszerre épül, akkor objektívnek nevezzük azokat a tulajdonságokat és relációkat amelyek a (definíciatlan) axiómákból logikai úton levezethetők.” A fenti két meghatározással kiegészítve LEIBNIZ kijelentése már értelmezhető.

Megjegyezzük továbbá, hogy a dolgozat lektorának véleménye szerint a fenti definíció helyett a következő megfogalmazás alkalmazandó:

Az \mathfrak{K} halmazt akkor nevezzük a T transzformációra invariánsnak, ha

1. $\forall r \in \mathfrak{K} : T(r) \in \mathfrak{K}$
2. Ha $s \in \mathfrak{K}$ akkor $\exists r \in \mathfrak{K} : T(r) = s$.

A definícióval analóg módon az $F(x, y)$ skalármező invariáns a C_n csoportra nézve, ha

$$\forall u_k \in C_n, \forall (x, y) \in \mathfrak{R}^2 : F(u_k(x, y)) = F(x, y).$$

A $C(F)$ számot F szimmetria csoportja rendjének nevezzük. A D_m tenzormező invariáns a C_n csoportra nézve, ha az $u_k(D_m)$ ($u_k \in C_n$) tenzormező koordináta-függvényei az eredetiekkel rendre megegyeznek. (Az u_k transzformáció végrehajtásának módját pl. DOMOKOS (1989) tárgyalja.)

2.5 Definíció. Az n -edfokú, homogén polinomokat $H^n()$ -nel jelöljük, a zárójelben a polinom változóit tüntetjük fel. (A változók száma nem rögzített.) Ha nem tüntetünk fel változót, akkor $H^n = H^n(x, y)$. Ezek szerint

$$H^n = H^n(x, y) = \sum_{i=0}^n a_i x^i y^{n-i}$$

2.6 Definíció. Az n -edfokú polinomokat $P^n(\cdot)$ -nel jelöljük, a zárójelben a polinom változóit tüntetjük fel. Ha nem tüntetünk fel változót, akkor $P^n = P^n(x, y)$. Ezek szerint

$$P^n = P^n(x, y) = \sum_{i=0}^n H^i = \sum_{i=0}^n \sum_{j=0}^i a_{i,j} x^i y^{i-j}$$

Térjünk át ezután hengerkoordinátákra a következő transzformációval:

$$(2.1) \quad \begin{aligned} x &= r \cos(\varphi) \\ y &= r \sin(\varphi). \end{aligned}$$

Az új koordináta-rendszerben

$$(2.2) \quad H^n(r \cos(\varphi), r \sin(\varphi)) = H_n(r, \varphi) = H^n(r) H^n(\cos(\varphi), \sin(\varphi)),$$

illetve

$$(2.3) \quad \begin{aligned} P^n(r \cos(\varphi), r \sin(\varphi)) &= P_n(r, \varphi) = \\ &= \sum_{i=0}^n H_i(r, \varphi) = \sum_{i=0}^n H^i(r) H^i(\cos(\varphi), \sin(\varphi)). \end{aligned}$$

Végezzük el ezután a H_n függvények *Fourier sorfejtését* a φ változó szerint:

$$(2.4) \quad H_n(r, \varphi) = r^n \sum_{i=0}^{\infty} (b_i \cos(i\varphi) + c_i \sin(i\varphi)) = r^n \sum_{i=0}^{\infty} d_i \cos(i\varphi - \varphi_i).$$

Ezen általános alak felírása után célunk, hogy az i index által felvehető értékek halmazát szűkítsük. Számítsuk ki a k -adik *Fourier-együttható* értékét

$$(2.5) \quad \begin{aligned} b_k &= \int_0^{2\pi} \sum_{i=0}^n a_i x^i y^{n-i} \cos(k\varphi) d\varphi = \\ &= r^n \int_0^{2\pi} \sum_{i=0}^n a_i \left(\frac{e^{j\varphi} + e^{-j\varphi}}{2} \right)^i \left(\frac{e^{j\varphi} - e^{-j\varphi}}{2} \right)^{n-i} \left(\frac{e^{jk\varphi} + e^{-jk\varphi}}{2} \right) d\varphi, \end{aligned}$$

ahol $j = \sqrt{-1}$.

Ha $k > n$, akkor a fenti integrál zérus, mivel a szorzás végrehajtása után nem keletkezik zérus kitevőjű tag a szummában. Ugyanez érvényes a c_k *Fourier-együtthatóra* is. A (2.5) összefüggés alapján tehát megállapíthatjuk, hogy a (2.4) egyenletben $i \leq n$. További szűkítést érhetünk el, ha figyelembe vesszük, hogy páros n esetén H^n C_2 szimmetriával rendelkezik, azaz $H^n(-x, -y) = H^n(x, y)$, páratlan

n esetén pedig $H^n(-x, -y) = -H^n(x, y)$. Ezek után esetünkben így alakul a (2.4) általános Fourier sor:

$$(2.6) \quad H_n(r, \varphi) = r^n \sum_{i=0}^{[n/2]} a_i \cos((2i + (n \bmod 2))\varphi - \varphi_i),$$

ahol $(n \bmod 2) = 1$ ha n páratlan és $(n \bmod 2) = 0$ ha n páros. Ezek után írjuk fel a $P_n(r, \varphi)$ függvényt (2.3) szerint:

$$(2.7) \quad P_n(r, \varphi) = \sum_{i=0}^n H_i(r, \varphi) = \sum_{i=0}^n r^i \sum_{j=0}^{[i/2]} a_{i,j} \cos((2j + (i \bmod 2))\varphi - \varphi_{i,j}).$$

Rendezzük át a kettős összegzést, mégpedig oly módon, hogy kigyűjtjük az azonos periódusú függvényeket:

$$(2.8) \quad P_n(r, \varphi) = \sum_{i=0}^n \sum_{j=[i/2]}^{[n/2]} u_{i,j}(H^k(r) \cos(i\varphi)),$$

ahol $u_{i,j}$ egy tetszőleges, origó körüli síkbeli forgatást jelöl, és $k = 2j + i \bmod 2$.

2.7 Definíció. Legyen

$$M^n(x, y) = x^n - \binom{n}{2} y^2 x^{n-2} + \binom{n}{4} y^4 x^{n-4} \dots$$

illetve

$$M_n(r, \varphi) = M^n(r \cos(\varphi), r \sin(\varphi)) = r^n \cos(n\varphi).$$

Az M^n függvényeket n -edrendű magpolinomoknak fogjuk nevezni. Legyen továbbá

$$K^n = H^n(x^2 + y^2)$$

illetve

$$(2.9) \quad K_n = K^n(r \cos(\varphi), r \sin(\varphi)) = H^n(r^2).$$

A K^n függvényeket (amelyek speciális, $2n$ -ed fokú homogén polinomok) n -edrendű körpolinomoknak fogjuk nevezni.

A (2.8) egyenlet és a (2.9) összefüggés szerint igaz a következő

2.1 LEMMA. Igaz a következő egyenlőség:

$$(2.10) \quad P^n(x, y) = \sum_{i=0}^n \sum_{j=0}^{[(n-i)/2]} u_{i,j}(K^j M^i).$$

A 2.1 lemmában megadott felbontás, mint látni fogjuk, sok szempontból hasonlít a *Fourier-sorra*. Érdekessége, hogy a bázisfüggvényeknek megfelelő M^n magpolinomok együtthatói a *Taylor-sor* együtthatóiból határozhatók meg. (Erre majd az alkalmazások között példát is mutatunk.)

Talán érdekes megemlítenünk, hogy a síkbeli *Laplace-differenciálegyenlet*

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

kapcsán fogalmazta meg DIRICHLET az alábbi problémát (SZŐKEFALVI-NAGY, 1977): Keressünk egy olyan $u(x, y)$ függvényt, amely egy síkbeli tartomány belsejében kielégíti a *Laplace-differenciálegyenletet* és a tartomány határa felé közeledve megadott határérték felé tart. Ha a tartomány egy kör belseje, akkor az általunk bevezetett jelölésekkel a

$$u(x, y) = \sum_{i=0}^n c_i u_i(M^i) \quad (n = 1, 2, \dots)$$

függvények ennek a problémának megoldásai. Az $u(x, y)$ függvényeket SZÁSZ (1951) harmonikus polinomoknak nevezi. Az említett probléma potenciálelméleti eredetű.

2.2 LEMMA. Ha $z = F(x, y)$ az origóban analitikus, akkor ezen pont pozitív környezetében

$$(2.11) \quad F(x, y) = \lim_{n \rightarrow \infty} \sum_{i=0}^n \left(\sum_{j=0}^{[(n-i)/2]} u_{i,j}(K^j M^i) \right)$$

alakban előállítható.

Bizonyítás. Ha $F(x, y)$ analitikus, akkor *Taylor-sora* konvergens, vagyis a konvergencia tartományon belül

$$(2.12) \quad F(x, y) = \lim_{n \rightarrow \infty} \sum_{i=0}^n H^n(x, y).$$

A 2.1 lemma és a 2.6 definíció alapján azonban a fenti függvénysor részösszege minden n -re megegyezik a 2.2 lemmában megadott függvénysor részösszegével, ezért a (2.11) függvénysor is konvergens, és előállítja az $F(x, y)$ analitikus függvényt. Q.e.d.

2.8 Definíció. $C(P^n) = m$ esetén a P^n polinomot $P_{(m)}^n$ módon fogjuk jelölni.

2.1 TÉTEL. Igaz a következő egyenlőség:

$$(2.13) \quad P_{(m)}^n(x, y) = \sum_{i=0}^{[n/m]} \sum_{j=0}^{[(n-im)/2]} u_{i,j}(K^j M^{im}).$$

Bizonyítás. A 2.1 lemmában megadott alakból következik, hogy $C(P^n)$ az összegben szereplő nem-zérus együttthatójú magpolinomok rendjének legnagyobb közös osztója. Ezek szerint tehát, ha $C(P^n) = m$, akkor csak im rendű magpolinomok fordulhatnak elő az összegben. Q.e.d.

A 2.1 tétel alapján kiszámítható a $\{P_{(m)}^n\}$ halmaz dimenziója, vagyis a független együttthatók száma.

2.2 TÉTEL.

$$(2.14) \quad \dim\{P_{(m)}^n\} = [n/2] + 1 + 2 \sum_{i=1}^{[n/m]} ((n - im)/2 + 1)$$

Bizonyítás. Vegyük figyelembe, hogy

$$\dim\{u_{i,j}(K^j M^{im})\} = 2,$$

ha $i \neq 0$, hiszen K^j és $u_{i,j}$ egyaránt egy-egy szabad paramétert jelent. Az $i = 0$ esetben a (2.9) összefüggés szerint $M^i = 1$, vagyis $C(M^i) = 0$. Ekkor tehát $\forall u_{i,j} : u_{i,j}(M^0) = M^0$, vagyis az $u_{i,j}$ elforgatás nem növeli a dimenziók számát:

$$\dim\{u_{0,j}(K^j M^0)\} = 1.$$

A fentiek, továbbá a 2.1 tételben megadott alak felhasználásával a 2.2 tételben kimondott egyenlőséghez jutunk. Néhány $\dim(P_{(m)}^n)$ értéket mutat az 1. táblázat.

A fenti táblázat alapján már sejthető a következő

2.3 TÉTEL.

$$(2.15) \quad \lim_{n \rightarrow \infty} \frac{\dim\{P_{(m)}^n\}}{n^2} = \frac{1}{2m}$$

Bizonyítás. Legyen

$$s_{max}^0(n, m) = \frac{n}{2} + 1 + \sum_{i=1}^{[n/m]} (n - im + 2)$$

$$s_{min}^0(n, m) = \frac{n}{2} + \frac{1}{2} + \sum_{i=1}^{[n/m]} (n - im + 1).$$

$m \rightarrow$	1	2	3	4	5	6	7	8	9	10
$n \downarrow$										
1	3	—	—	—	—	—	—	—	—	—
2	6	4	—	—	—	—	—	—	—	—
3	10	4	4	—	—	—	—	—	—	—
4	15	9	5	5	—	—	—	—	—	—
5	21	9	7	5	5	—	—	—	—	—
6	28	16	10	8	6	6	—	—	—	—
7	36	16	12	8	8	6	6	—	—	—
8	45	25	15	13	9	9	7	7	—	—
9	55	25	19	13	11	9	9	7	7	—
10	66	36	22	18	14	12	10	10	8	8

1. táblázat: Az n -ed fokú, m -ed rendű szimmetriával rendelkező polinomok dimenziója.

Legyen továbbá

$$(2.16) \quad s_{\max}(n, m) = \frac{n^2 + 4n + 2m(n+1)}{2m}$$

$$(2.17) \quad s_{\min}(n, m) = \frac{n^2 + 2n - m(2n+1)}{2m}.$$

Megállapíthatjuk, hogy

$$(2.18) \quad \forall n, m : s_{\min} \leq s_{\min}^0 \leq \dim\{P_{(m)}^n\} \leq s_{\max}^0 \leq s_{\max}.$$

Mivel

$$(2.19) \quad \lim_{n \rightarrow \infty} \frac{s_{\max}}{n^2} = \lim_{n \rightarrow \infty} \frac{s_{\min}}{n^2} = \frac{1}{2m},$$

ezért a (2.18) egyenlőtlenségek alapján

$$(2.20) \quad \lim_{n \rightarrow \infty} \frac{\dim\{P_{(m)}^n\}}{n^2} = \frac{1}{2m}.$$

Q.e.d.

A 2.3 tétel alapján triviális az alábbi

2.3 LEMMA.

$$(2.21) \quad \lim_{n \rightarrow \infty} \frac{\dim\{P_{(k)}^n\}}{\dim\{P_{(l)}^n\}} = \frac{l}{k}$$

A továbbiakban a 2.1 tételhez mondunk ki néhány korolláriumot, amelyek elsősorban az alkalmazások során lesznek érdekesek.

2.1 KOROLLÁRIUM A (2.1) TÉTELHEZ. Ha $H^n(x, y)$ nem azonosan zérus, akkor $C(H^n(x, y))$ lehetséges értékei a $2k + (n \bmod 2)$ alakú számok, ahol $0 \leq k \leq n/2$.

Bizonyítás. A (2.6) egyenlet és a (2.9) összefüggés szerint

$$H^n(x, y) = \sum_{i=0}^{[n/2]} u_i (M^k K^{(n-k)/2}),$$

ahol u_i egy síkbeli forgatást jelöl és $k = 2i + (n \bmod 2)$. A fenti alakból, továbbá a $C(M^n) = n$ egyenlőségből következik, hogy a korolláriumban megjelölt szimmetriák létrejöhetnek, és csak azok jöhetnek létre. Q.e.d.

2.2 KOROLLÁRIUM A (2.1) TÉTELHEZ. $C(P^n(x, y)) \leq n$ és n a legkisebb felső korlát.

Bizonyítás. A 2.1 tétel szerint $C(P^n) = m$ esetén M^{im} -nek ($i = 1, 2, \dots$) szerepelnie kell a kettős összegben. Mivel M^{im} (im)-ed fokú polinom, ezért $im \leq n$, amiből a korollárium első állítása következik. Legyen ezután $P^n = M^n$, ekkor nyilván $C(P^n) = n$. Ezzel a korollárium állításának második részét is beláttuk. Q.e.d.

2.9 Definíció. A kétdimenziós $z = F(x, y)$ skalármező n -edrendű Taylor-közelítését a $P(0, 0)$ pontban $F^n(x, y)$ fogja jelölni. Az F^n Taylor-közelítést akkor nevezzük szimmetria-invariánsnak, ha $C(F) = C(F^n)$.

2.3 KOROLLÁRIUM A (2.1) TÉTELHEZ.

$$(2.22) \quad C(F^n) = C(F) \quad \text{szükséges feltétele, hogy} \quad n \geq C(F).$$

Bizonyítás. A 2.2 korollárium alapján $C(F^n) \leq n$, mivel F^n n -ed fokú polinom, ezért $C(F) = C(F^n)$ esetén $C(F) \leq n$ is fennáll, ami a korollárium állításával egyenértékű. Q.e.d.

2.4 KOROLLÁRIUM A 2.1 ÉS 2.2 LEMMÁHOZ. Ha F az origóban analitikus, akkor

$$(2.23) \quad C(F^n) = kC(F) \quad (n, k = 0, 1, 2, \dots)$$

Bizonyítás. A 2.1 és 2.2 lemma szerint mind az F , mind az F^n függvényre igaz, hogy szimmetriacsoportjuk rendje megegyezik az összegzésben szereplő magpolinomok rendjének legnagyobb közös osztójával. Mivel minden, az F^n függvényben szereplő magpolinom egyben az F függvényben is szerepel, ezért $C(F^n) = kC(F)$. Q.e.d.

2.2 Általános konzervatív tenzorok

2.10 Definíció. A D_m m -edrendű, konzervatív tenzormező n -edrendű Taylor-közelítését D_m^n -nel fogjuk jelölni.

A D_m^n közelítés a D_m tenzormező 2^m darab skalár koordinátájának n -edrendű Taylor-közelítéséből áll.

2.4 TÉTEL. A $C(D_m^n) = C(D_m)$ egyenlőség szükséges feltétele, hogy $n \geq C(D_m) - m$.

Bizonyítás. Azok az F függvények, amelyek m -edrendű deriváltja D_m , egymástól egy $(m-1)$ -edrendű polinomban különbözhetnek. Válasszuk ki közülük azt az F_0 függvényt, amelynél az F_0^{m-1} közelítő Taylor-polinom minden együtthatója zérus, vagyis a Taylor-sor az m -ed fokú tagokkal kezdődik. Belátjuk, hogy erre teljesül a következő

2.4 LEMMA.

$$(2.24) \quad C(F_0) = C(D_m)$$

Bizonyítás. Legyen $u_k \in C_{C(F_0)}$ és forgassuk el az $[xy]$ koordináta-rendszert u_k -val. Ekkor a 2.4 definíció alapján F_0 önmagába transzformálódik, tehát deriváltjai azonosak az elforgatott helyzetben is. Ezek szerint tehát $C(D_m) = jC(F_0)$ ($j = 1, 2, \dots$) kell, hogy legyen.

Ugyancsak a 2.4 definícióra hivatkozva megállapíthatjuk, hogy a $v_k \in C_{C(D_m)}$ transzformációval elforgatott rendszerben D_m koordináta-függvényei megegyeznek az eredetiekkel. Ezen függvények alapján viszont — az integrálási polinomoktól eltekintve — F_0 ill. F_0 elforgatottja egyértelműen meghatározható. Mivel F_0 $(m-1)$ -edik Taylor közelítő polinomja azonosan zérus és ezért ugyanez áll F_0 elforgatottjára

is, ezért a két függvény különbsége, amely legfeljebb $(m-1)$ -ed fokú polinom, azonosan zérus. Ezek szerint tehát $C(F_0) = kC(D_m)$ ($k = 1, 2, \dots$).

Összefoglalva tehát:

$$(2.25) \quad \begin{aligned} (i) \quad & C(D_m) = jC(F_0) \\ (ii) \quad & C(D_m) = \frac{1}{k}C(F_0) \end{aligned} \quad (j, k = 1, 2, \dots)$$

A fentiek alapján nyilván $j = k = 1$, és ezzel a lemma állítását beláttuk. Q.e.d.

Közbevetőleg belátjuk azt is, hogy igaz a következő

2.5 LEMMA. *Ha D_m nem-zérus, konstans tenzormező, akkor $C(D_m) = 2k + (m \bmod 2)$, $(0 \leq k \leq n/2)$.*

Bizonyítás. Ha $D_m(x, y) = \text{konstans} \neq 0$, akkor F_0 homogén, m -edfokú, nem-zérus polinom. A 2.4 lemma alapján $C(D_m) = C(F_0)$, a 2.1 korollárium és a bizonyítás első megállapítása alapján pedig $C(F_0) = 2k + (m \bmod 2)$, $(0 \leq k \leq n/2)$. Ezzel a lemma állítását beláttuk. Q.e.d.

Ezek után folytatjuk a 2.4 tétel bizonyítását.

Vegyük figyelembe, hogy az n -edrendű D_m^n Taylor-közelítés maga is m -edrendű tenzormező, hiszen azonos az F_0 felület $(n+m)$ -edrendű Taylor-közelítésének m -edrendű kovariáns deriváltjával. Ily módon a (36) egyenletet a csonkolt Taylor-sorra is alkalmazhatjuk, hiszen F_0^{n+m} -re is teljesül, hogy az $(m-1)$ -ed fokú Taylor közelítő polinomja azonosan zérus, így:

$$(2.26) \quad C(D_m^n) = C(F_0^{n+m}).$$

A 2.3 korollárium alapján:

$$(2.27) \quad C(F_0^{n+m}) = C(F_0) \quad \text{csak akkor, ha} \quad n+m \geq C(F_0).$$

A (2.24) és (2.26) egyenletet a (2.27) egyenletbe helyettesítve a

$$(2.28) \quad C(D_m^n) = C(D_m) \quad \text{csak akkor, ha} \quad n \geq C(D_m) - m.$$

összefüggésre jutunk. Q.e.d.

Megjegyezzük, hogy a 2.4 tétel nem zárja ki, hogy $C(D_m) - m$ negatív értéket vegyen fel. Ebben az esetben biztosak lehetünk benne, hogy a szükséges és elégséges feltétel nem azonos. Van rá azonban mód, hogy a negatív értékeket kizárjuk a 2.4 tétel feltételéből. Ennek igazolásához a 2.5 lemmára támaszkodunk, ahol beláttuk, hogy egy nem-zérus, konstans tenzormező rendjének paritása megegyezik szimmetria-csoportja rendjének paritásával. Ezek szerint tehát, ha

$$[C(D_m) \bmod 2] = [m \bmod 2] \quad (C(D_m) \leq m),$$

akkor a D_m tenzormező konstans (0-adrendű) *Taylor-sorának* is lehetnek $C(D_m)$ -edrendű harmonikus komponensei, egyébként ($D_m^0 \neq 0$ esetén) ez nem lehetséges, tehát legalább a lineáris (elsőrendű) *Taylor-tagokig* el kell menni. Ezek szerint $C(D_m) \leq m$ esetén a szükséges *Taylor-tagok* rendje 0 vagy 1.

A fenti állítást úgy is megfogalmazhatjuk, hogy ha a tenzormező rendjének és a szimmetria rendjének paritása megegyezik, akkor csak $n \geq 0$ állítható, ellenkező esetben ($D_m^0 \neq 0$ esetén) $n \geq 1$ is teljesül.

A 2.4 tétel állítását ezzel a megfontolással kiegészítve a 2. táblázat illusztrálja:

$m \rightarrow$	0	1	2	3	4	5	6	7	8	9
$C(D_m) \downarrow$										
0	0	0	0	0	0	0	0	0	0	0
1	1	0	1	0	1	0	1	0	1	0
2	2	1	0	1	0	1	0	1	0	1
3	3	2	1	0	1	0	1	0	1	0
4	4	3	2	1	0	1	0	1	0	1
5	5	4	3	2	1	0	1	0	1	0
6	6	5	4	3	2	1	0	1	0	1
7	7	6	5	4	3	2	1	0	1	0
8	8	7	6	5	4	3	2	1	0	1
9	9	8	7	6	5	4	3	2	1	0

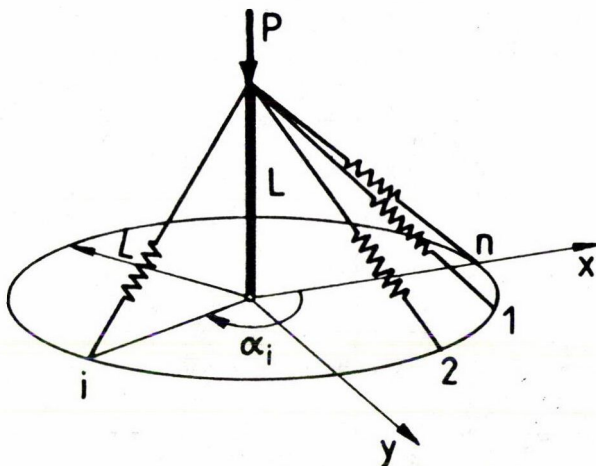
2. táblázat: A Taylor-sorfejtés szükséges rendje

3. Példák

3.1. A potenciális energia: $m = 0$

A V potenciális energia skalár mennyiség, vagyis nulladrendű tenzor. A 2.3 korollárium állítása szerint ekkor legalább a $C(V)$ -edrendű tagokig kell menni a

Taylor-sorfejtésben, hogy a szimmetria rendje ne változzon. Illusztrációként GÁSPÁR (1977) munkáját említjük, melyben egy n ($n \geq 3$) szimmetrikusan elhelyezett lineáris rugóval megtámasztott merev rúd viselkedését vizsgálja. A szerkezetet az 1. ábra illusztrálja.



1. ábra: Az n rugóval megtámasztott rúd

GÁSPÁR megállapítja, hogy a szerkezet potenciális energiájának ötödrendű Taylor-közelítése a

$$(3.1) \quad V^5 = KL^2 \left(\frac{1}{8}r^2 + \frac{1}{8}C_3r^3 + \frac{15}{512}r^4 + \frac{5}{64}C_4r^4 + \frac{7}{128}C_5r^5 \right) + PL \left(-\frac{1}{2}r^2 - \frac{1}{8}r^4 \right) - \frac{1}{4}KL^2rr_0 \cos(\varphi - \varphi_0)$$

formában írható, ahol

$$\begin{aligned} K, L, P & \text{ konstansok} \\ C_3 &= \begin{cases} \frac{1}{4} \cos(3\varphi) & \text{ha } n = 3 \\ 0 & \text{ha } n \neq 3 \end{cases} \\ C_4 &= \begin{cases} \frac{1}{8} \cos(4\varphi) & \text{ha } n = 4 \\ 0 & \text{ha } n \neq 4 \end{cases} \\ C_5 &= \begin{cases} \frac{5}{16} \cos(3\varphi) & \text{ha } n = 3 \\ \frac{1}{16} \cos(5\varphi) & \text{ha } n = 5 \\ 0 & \text{egyébként} \end{cases} \\ r_0, \varphi_0 & \text{ tökéletlenségek.} \end{aligned}$$

A szerkezet n -szeres forgásszimmetriája miatt n rugó esetén a tökéletes ($r_0 = 0$) szerkezeten

$$(3.2) \quad C(V) = kn \quad (k = 1, 2, \dots).$$

A (3.1) képletből viszont leolvasható, hogy $r_0 = 0$ -ra, $n = 3, 4, 5$ esetén

$$(3.3) \quad C(V^n) = n.$$

Mivel a 2.4 korollárium alapján $C(V^n) = jC(V)$ ($j = 1, 2, \dots$), ezért $k = 1$ teljesül a (3.2) egyenletben. A fenti egyenletbe behelyettesítve a (3.2) egyenletet $k = 1$ helyettesítéssel, a

$$(3.4) \quad C(V^{C(V)}) = C(V)$$

összefüggésre jutunk, amely szerint a 2.3 korolláriumban megfogalmazott szükséges feltétel itt egyben elégséges is.

3.2. A merevségi operátor: $m = 1$

A K merevségi operátor az u elmozdulás-vektorhoz a q tehervektort rendeli hozzá:

$$(3.5) \quad K(u) = q.$$

A K operátor tehát egy vektormezőt, vagy másképp kifejezve, egy elsőrendű tenzormezőt definiál, amely nem más, mint a potenciális energia gradiense, tehát a tenzormező konzervatív. A 2.4 tétel szerint ekkor legalább a $(C(K) - 1)$ -edrendű tagokig kell elmennünk a *Taylor-sorfejtésben*, hogy a szimmetria rendje ne változzék. Illusztrációként a kikötött árbócok korlátozottan nagy elmozdulásainak numerikus számítását említjük.

Korlátozottan nagy elmozdulások esetén az egyensúlyi feltételt kifejező nemlineáris egyenletrendszer iterációs megoldására a módosított *Newton-Raphson módszer* (GÁSPÁR, 1983) szokták alkalmazni.

Ennél az eljárásnál a kiindulási helyzet linearizált K^1 merevségi operátorát alkalmazzák az egész terhelési folyamat során.

A lineáris közelítésben minden koordináta irányú erő csak a saját irányába idéz elő elmozdulást, amiből következik, hogy a K^1 mátrix diagonális. A további vizsgálatoknál csak azt fogjuk feltenni, hogy a vízszintes és a függőleges erők között nincs kapcsolat, vagyis hogy a $K(x, y, z)$ operátor egy $K_h(x, y)$ és egy $K_v(z)$ operátorra bontható.

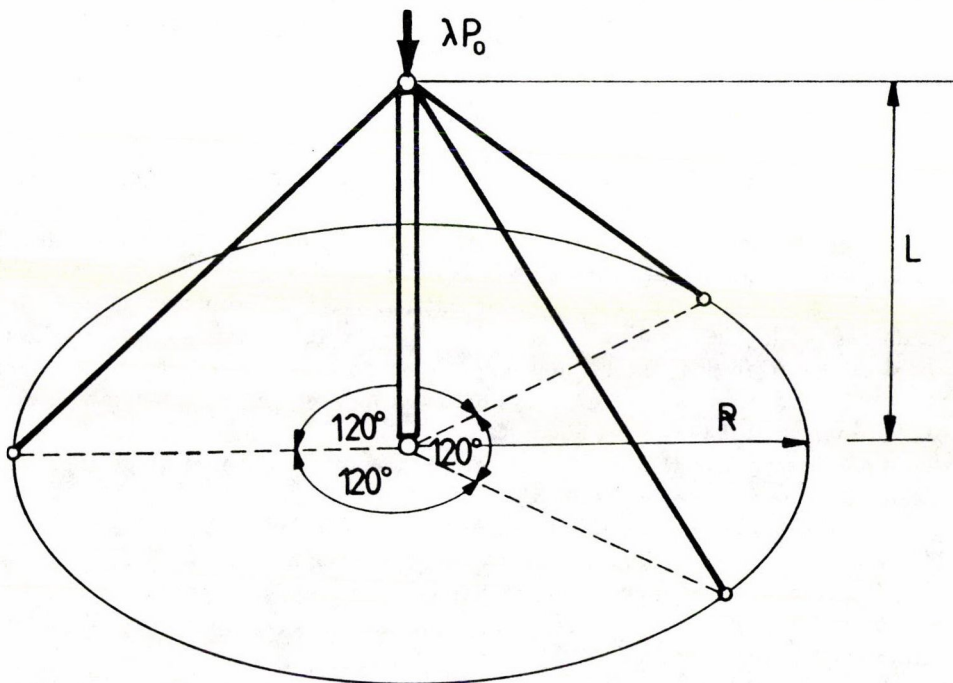
Kikötött árbócok esetében a módosított *Newton-Raphson módszer* közvetlen alkalmazása több szempontból is hátrányos lenne. Ennek egyik oka, hogy $n > 2$

kötél esetén a tökéletes szerkezet K_h^1 merevségi operátora C_0 szimmetriát mutat, tehát elfedi a szerkezet diszkrét forgási szimmetriáját. Másrészt a K_v^1 egydimenziós vektormező C_2 szimmetriát mutat, tehát a szerkezet látszólag húzásra és nyomásra egyformán viselkedik, a valóságban azonban ilyen szimmetria nincs, tehát $C(K_v) = 1$. Mindkét tény arra mutat, hogy a lineáris operátor segítségével csak meglehetősen torz képünk alakulhatna ki a szerkezet viselkedéséről. Ezért, megtarva a módosított *Newton-Raphson módszer* alapgondolatát, az eljárást azzal a követelménnyel egészítjük ki, hogy a kezdeti (terheletlen) állapotban számított, és a teljes terhelés alatt alkalmazott csnokolt K_h^n és K_v^m merevségi operátoroktól megkívánjuk, hogy szimmetria-invariánsak legyenek, vagyis $C(K_h^n) = C(K_h)$ és $C(K_v^m) = C(K_v)$ teljesüljön.

Példaként egy 3 kötéllal merevített árbócot mutatunk be. Ekkor

$$(3.6) \quad \begin{array}{ll} (i) & C(K_h) = 3 \\ (ii) & C(K_v) = 1 \end{array}$$

A (3.6) (ii) egyenlőség természetesen $k > 0$ kötélen mindig teljesül. A szerkezet elrendezését a 2. ábra mutatja ($\lambda < 0$):



2. ábra: A vizsgált áróc

Számszerű adatok:

$$R = 2000 \text{ cm}$$

$$L = 1000 \text{ cm}$$

$$EA_{\text{rúd}} = 2.10 \times 10^6 \text{ kN}$$

$$EA_{\text{kötél}} = 2.92 \times 10^4 \text{ kN}$$

$$EI_{\text{rúd}} = 1.75 \times 10^7 \text{ kNcm}^2$$

$$l_{\text{kötél}} = 2231.2 \text{ cm (nyúlásmentes hossz)}$$

$$P_0 = 1 \text{ kN}$$

A 2.4 tétel alapján megállapíthatjuk, hogy a K_h^n csnokolt sorfejtésben $n_{\text{szükséges}} = 2$. Ugyanezen tétel alapján a K_v^m csnokolt sorfejtésben $m_{\text{szükséges}} = 0$. Mivel $C(K_v^1) \neq C(K_v)$, ezért ez nyilván nem elégséges. Ekkor nyilván $m_{\text{szükséges}} \geq 2$. Megjegyezzük, hogy n és m szükséges értékei csak 3 kötélen esnek egybe. Feladatunk a K_h^2 és K_v^2 operátorok numerikus előállítására. A deriváltak értékeinek kiszámítására a differenciámódszert alkalmaztuk. A vizsgált „rácst” a 3. ábra mutatja. (A rácsoldalak méreteit nagyítottuk.)

Az i -edik rácspontban számított erőt F_i -vel jelöljük, komponenseit pedig az $F_{i,x}, F_{i,y}, F_{i,z}$ módon. A 9. rácspont azt a helyzetet jelöli, amikor a szerkezet $\lambda = 0$ mellett egyensúlyban van. Emiatt, valamint a 10. és 11. rácspontokon szimmetria miatt megállapíthatjuk, hogy

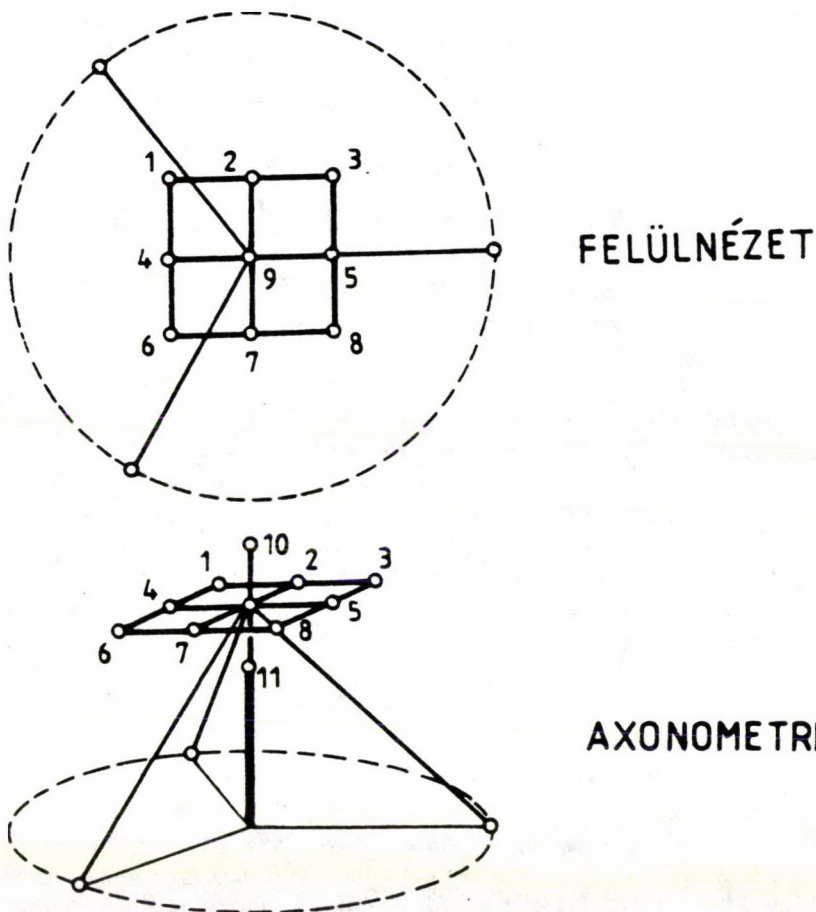
$$(3.7) \quad F_9 = 0, \quad F_{10,x} = F_{10,y} = 0, \quad F_{11,x} = F_{11,y} = 0.$$

Az egyszerűség kedvéért a K^2 operátorokat is mátrixos alakban foglaljuk össze, annak ellenére, hogy nem lineáris transzformációt definiálnak. Ez az operátor nem más, mint a potenciális energia gradiensének másodfokú *Taylor-sora*, amelynek elemei azonosak magának a potenciális energia harmadfokú *Taylor-sorának* elemeivel. Használjuk fel a 2.1 és 2.2 tételt annak eldöntésére, hogy az imént említett harmadfokú *Taylor-sor* egyes együtthatói között mi az összefüggés. A 2.2 tétel, illetve az 1. táblázat szerint $\dim(P_{(3)}^3) = 4$, tehát a 10 együttható közül mindössze 4 független. A (2.9) összefüggés és a 2.1 tétel szerint $P_{(3)}^3$ általános alakja a következő

$$(3.8) \quad P_{(3)}^3 = ax^3 - 3axy^2 - by^3 + 3bx^2y + cx^2 + cy^2 + d.$$

(Megjegyezzük, hogy az $u_{i,j}$ elforgatást a $b \sin(3\varphi)$ függvény hozzáadásával helyettesítettük.) A (3.8) egyenlet és a megelőző megfontolások alapján a kétdimenziós K_h^2 operátort a

$$(3.9) \quad K_h^2 = \begin{bmatrix} 3a & -3a & 6b & 2c & 0 \\ 3b & -3b & -6a & 0 & 2c \end{bmatrix}$$



3. ábra: A differencia-rács

alakban írhatjuk, ahol az oszlopokban rendre a $u_x^2, u_y^2, u_x u_y, u_x, u_y$ változók együtt-hatói állnak. Megállapíthatjuk, hogy a szerkezet tengelyes szimmetriája miatt $b = 0$, tehát csak a és c értékét kell meghatároznunk. Bevezetve a

$$(3.10) \quad u_h^2 = \begin{bmatrix} u_x^2 \\ u_y^2 \\ u_x u_y \\ u_x \\ u_y \end{bmatrix}, \quad q_h = \begin{bmatrix} F_x \\ F_y \end{bmatrix}$$

jelölést, az erő-elmozdulás összefüggés a

$$(3.11) \quad K_h^2 u_h^2 = q_h$$

formában írható. A K_h^2 mátrix k_{ij} elemeit a 3. ábra alapján a következő módon fejezhetjük ki:

$$\begin{aligned} k_{11} &= \frac{F_{5,x} + F_{4,x}}{2\Delta x^2} & k_{21} &= \frac{F_{5,y} + F_{4,y}}{2\Delta x^2} \\ k_{12} &= \frac{F_{2,x} + F_{7,x}}{2\Delta y^2} & k_{22} &= \frac{F_{2,y} + F_{7,y}}{2\Delta y^2} \\ k_{13} &= \frac{(F_{3,x} + F_{6,x}) - (F_{1,x} + F_{8,x})}{4\Delta x \Delta y} & k_{23} &= \frac{(F_{3,y} + F_{6,y}) - (F_{1,y} + F_{8,y})}{4\Delta x \Delta y} \\ k_{14} &= \frac{F_{5,x} - F_{4,x}}{2\Delta x} & k_{24} &= \frac{F_{5,y} - F_{4,y}}{2\Delta x} \\ k_{15} &= \frac{F_{2,x} - F_{7,x}}{2\Delta y} & k_{25} &= \frac{F_{2,y} - F_{7,y}}{2\Delta y} \end{aligned}$$

Megjegyezzük, hogy a

$$\begin{bmatrix} k_{14} & k_{15} \\ k_{24} & k_{25} \end{bmatrix}$$

mátrix azonos az elsőrendű K_h^1 merevségi mátrixszal. A fenti megfontolások szerint elegendő a k_{11} és a k_{14} elemet meghatározni.

A K_v^2 operátornál lényegesen egyszerűbb a helyzet, hiszen ez csak egydimenziós, így összesen 2 komponense van:

$$(3.12) \quad K_v^2 = \begin{bmatrix} v_{11} & v_{12} \end{bmatrix} \leftarrow F_z$$

$$\begin{array}{cc} \uparrow & \uparrow \\ u_z^2 & u_z \end{array}$$

A konvenciót a (3.9) képlethez hasonlóan jelöltük, és az erő-elmozdulás összefüggés itt is a

$$(3.13) \quad u_v^2 = \begin{bmatrix} u_z^2 \\ u_z \end{bmatrix}, \quad q_v = F_z$$

jelölés bevezetése után a

$$(3.14) \quad K_v^2 u_v^2 = q_v$$

formában foglalható össze. A K_v^2 mátrix komponensei az előzőekhez hasonlóan számíthatók:

$$(3.15) \quad v_{11} = \frac{F_{10} + F_{11}}{\Delta z^2}, \quad v_{12} = \frac{F_{10} - F_{11}}{2\Delta z}.$$

A peremérték feladatok megoldása után az alábbi numerikus eredményekre jutottunk:

$$(3.16) \quad a = -3.5, \quad c = 20.8.$$

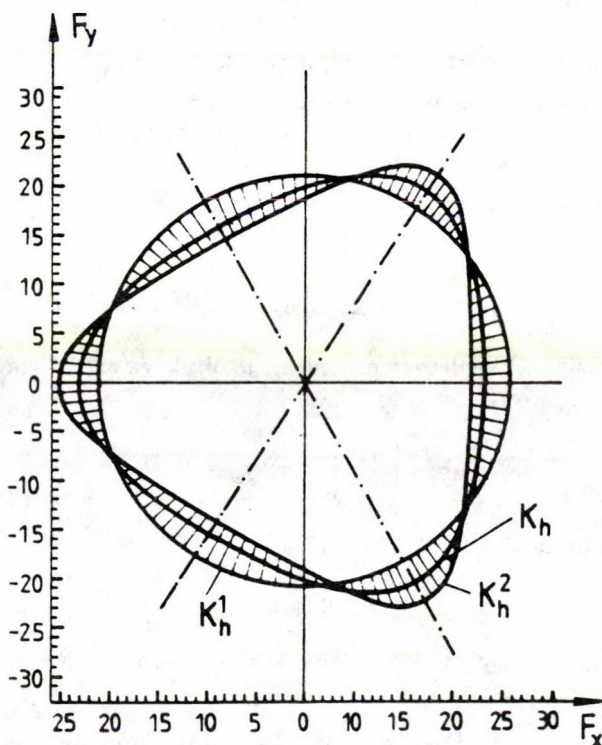
A (3.9) egyenlet szerint tehát

$$(3.17) \quad K_h^2 = \begin{bmatrix} -10.5 & 10.5 & 0 & 41.6 & 0 \\ 0 & 0 & 21.0 & 0 & 41.6 \end{bmatrix}$$

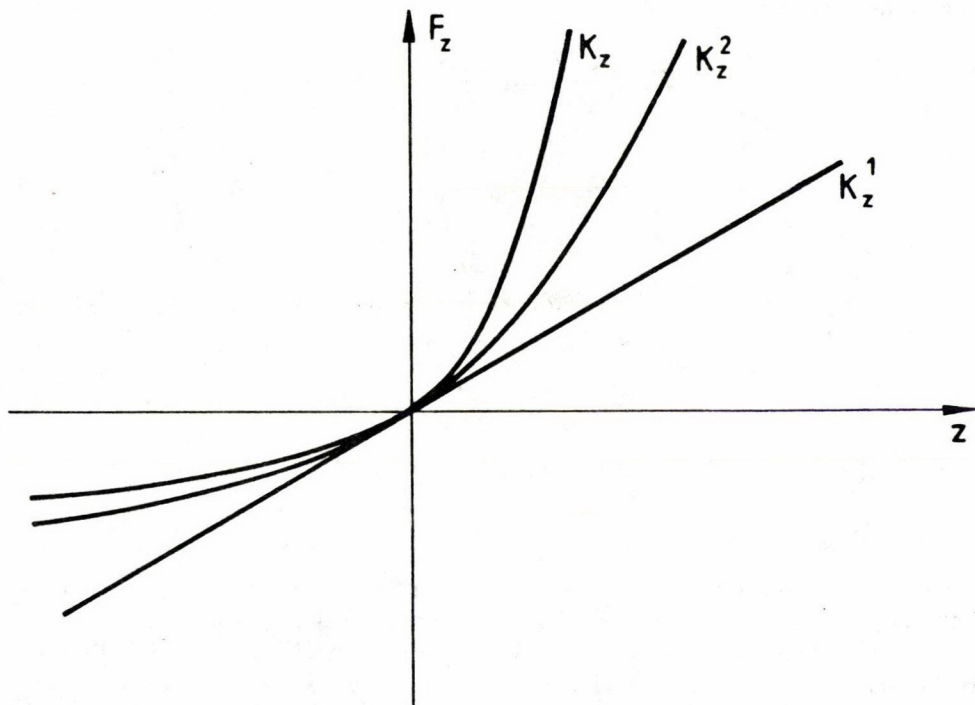
További numerikus eredmények alapján

$$K_v^2 = [12.3 \ 82.5]$$

Az operátorok birtokában már rajzolható erő-elmozdulás diagram. A vízszintes síkban egy $u_x^2 + u_y^2 = konstans$ körpályán mozgattuk a felső csomópontot. Kiszámítottuk az ehhez az elmozduláshoz rendelt erőket a K_h^1 , K_h^2 és K_h operátorok alkalmazásával. (Az utóbbi görbéhez 64 pontot számítottunk ki.) Az eredményt a 4. ábra mutatja. Láthatjuk, hogy a 2.4 tétel által meghatározott szükséges feltétel itt elégségesnek is bizonyult, hiszen a másodrendű operátorral számított görbe C_3 szimmetriát mutat. A függőleges irányban a teljes (egydimenziós) elmozduláshoz rendelt erőt ábrázoltuk az 5. ábrán.



4. ábra: A kiegyensúlyozatlan vízszintes erő diagramjai



5. ábra: A kiegyensúlyozatlan függőleges erő diagramjai

Megjegyezzük, hogy a 2.4 tétel által kimondott szükséges feltétel elégségességét illusztrálja hajlított rudakra (de természetesen nem elemi rúdelmélettel) BUZANO et al. (1985).

Megjegyezzük, hogy feszített árbócok merevségi operátorát vizsgálva HATHOUT et al. (1979) megállapítják, hogy a C_3 szimmetriával rendelkező szerkezet linearizált K^1 operátorára $C(K^1) = 0$ teljesül.

3.3. A feszültségi tenzor: $m = 2$

A feszültségi tenzor a vizsgált test minden pontjában egy irányhoz egy feszültségvektort rendel hozzá, tehát másodrendű tenzor. Illusztrációként a membránhéjak feszültségállapotát vizsgáljuk egy CSONKA (1981) által közölt példa alapján. CSONKA a vetületi \bar{N}_x , \bar{N}_{xy} , \bar{N}_y membránérőkkel dolgozik. Ezek az $F(x, y)$ feszültségfüggvény második parciális deriváltjaiként állnak elő, tehát az \bar{N}_2 membrántenzor eleget tesz a konzervatív tenzormező 2.1 definíciójának.

Szabályos háromszög alaprajzú, tehát C_3 szimmetriával rendelkező, egyenletesen megoszló teherrel terhelt héjakra CSONKA (1981, 5.1.(4), p.281) az alábbi

membránerőket közli:

$$\begin{aligned}
 \bar{N}_x &= \frac{\partial^2 F}{\partial y^2} = -\frac{\bar{p}_0 a_0}{h_0} (a_0 - x) \\
 \bar{N}_{xy} &= -\frac{\partial^2 F}{\partial y \partial x} = -\frac{\bar{p}_0 a_0}{h_0} y \\
 \bar{N}_y &= \frac{\partial^2 F}{\partial x^2} = -\frac{\bar{p}_0 a_0}{h_0} (a_0 + x).
 \end{aligned}
 \tag{3.18}$$

Mint látjuk, a vetületi feszítőerők kivétel nélkül az x és y változók lineáris függvényei, tehát

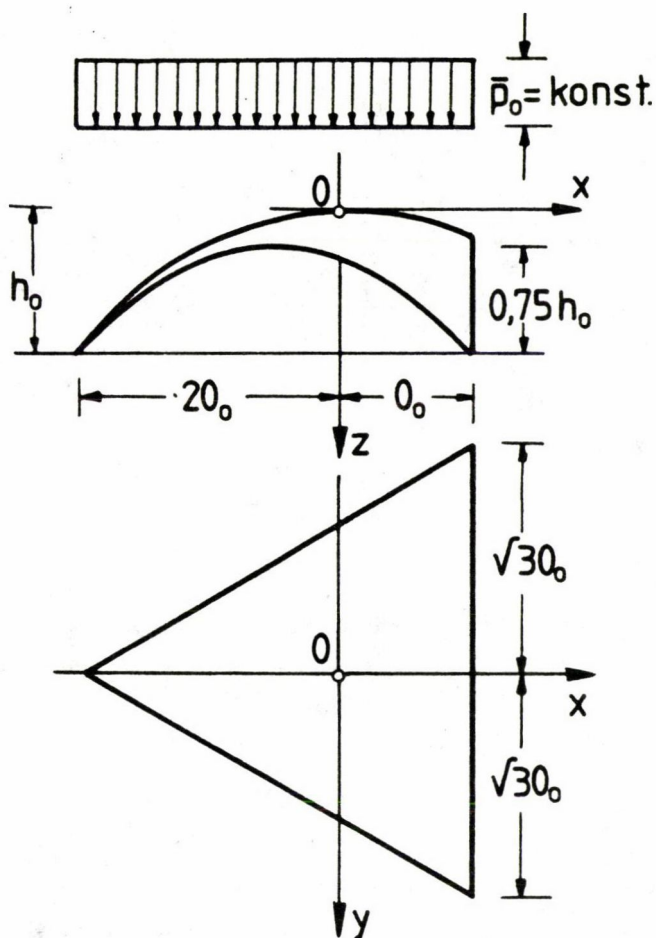
$$\bar{N}_2 = \bar{N}_2^1,
 \tag{3.19}$$

amiből természetesen

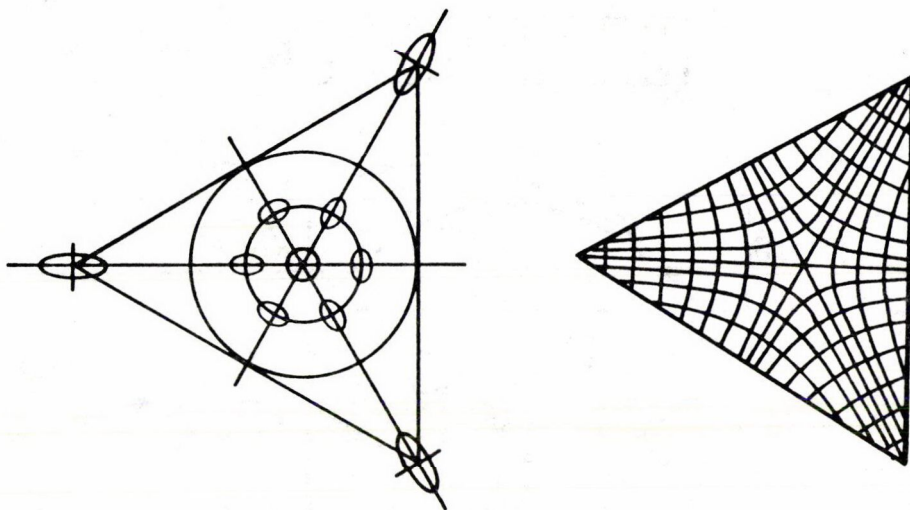
$$C(\bar{N}_2) = C(\bar{N}_2^1)
 \tag{3.20}$$

is következik. Mivel $C(\bar{N}_2) = 3$ és $m = 2$, ezért a 2.4 tétel szerint $n_{\text{szükséges}} = 1$. A (3.20) egyenlőség azt mutatja, hogy ez az érték egyben elégséges is.

CSONKA könyvében más szimmetrikus héjak is illusztrálják a 2.4 tétel állítását. Az \bar{N}_2 tenzormező C_3 szimmetriáját CSONKA könyvéből átvett ábrákkal próbáljuk szemléltetni. Elsőként a feladatot mutatjuk be (6/a ábra, CSONKA 5.1.2 ábrája alapján), majd a számítási eredményeket (6/b ábra, CSONKA 5.1.5-5.1.14. ábrái alapján).



6/a. ábra: A vizsgált héj



6/b. ábra: Membránfeszültségek c_3 szimmetriája

Köszönetnyilvánítás. A dolgozat megírása során folytatott konzultációért a szerző GÁSPÁR ZSOLTNak, MAKAI ENDRÉnek és GRIGOR ALEKSZANDROVICS MARGULISZnak mond köszönetet.

IRODALOM

- [BIRKHOFF, G. and MACLANE, S.] *A survey of modern algebra* (Macmillan, New York, 1965).
- [BUZANO, E., GEYMONAT, G. and POSTON, T.] "Post-buckling behaviour of a non-linearly hyperelastic thin rod with cross section invariant under the dihedral group D_n ", *Arch. for Rat. Mech.* **89** (1985), 307–388.
- [COXETER, H.S.M.] *Introduction to geometry* (John Wiley, New York, 1969).
- [CSONKA, P.] *Héjszerkezetek* (Akadémiai Kiadó, Budapest, 1981).
- [DIEUDONNÉ, J.A. and CARRELL, J.B.] *Invariant theory, Chap. 3,4.* (Academic Press, New York, London, 1971).
- [DOMOKOS, G.] "The equilibrium equations of membrane shells expressed in general surface coordinates", *Periodica Polytechnica T.U. Budapest/Arch.* **34** No. 1–2. (1989), 85–134.
- [GÁSPÁR, Zs.] "Buckling models for higher catastrophes", *J. Struct. Mech.* **5** (1977), 357–368.
- [GÁSPÁR, Zs.] *A katasztrófaelmélet alkalmazása a szerkezetek stabilitásvizsgálatában* (BME Mérnöki Továbbképző Intézet, Budapest, 1983), (Kézirat).
- [GÁSPÁR, Zs. and Tornyos, Á.] "Rúdszerkezetek állapotjellemzőinek számítása csomóponti koordinátarendszerben", *Magyar Építőipar* **3** (1986), 161–163.
- [HATHOUT, I.A. et al.] "Buckling of stayed columns", *ASCE*, vol. 105., No. ST9 (1979), pp. 1805–1822.
- [HEGEDŰS, I.] "Egyrétegű hengeres rácsok vizsgálata mátrixok segítségével", *Műszaki Tudomány* **56** (1978), 431–443.
- [KORN, G.A. and KORN, T.A.] *Mathematical Handbook* (McGraw-Hill, New York, 1968).
- [LEIBNIZ, G.W.] "Initia rerum Mathematicorum metaphysica", *Leibnizens Mathematische Schriften*, VII., p. 17, illetve V., p. 178, ed. Gerhardt C.J. (Berlin, 1848).

- [SCHWARZ, G.] "Smooth functions invariant under the action of a compact Lie group", *Topology* 14 (1975), 63–68..
- [SZÁSZ, P.] *A differenciál- és integrálszámítás elemei* (Tankönyvkiadó, Budapest, 1951), pp. 362.
- [SZŐKEFALVI-NAGY, B.] *Valós függvények és függvénysorok* (Tankönyvkiadó, Budapest, 1977), pp. 316.
- [TARNAI, T.] "Simultaneous static and kinematic indeterminacy of space trusses with cyclic symmetry", *Int. J. Solids Struc.* 16 (1980), 347–359.
- [THOM, R.] *Structural stability and morphogenesis* (Benjamin, New York, 1975).
- [TORNÝOS, Á.] "Dynamic analysis of space frameworks with cyclic symmetry", *Space Structures* 1 (1985), 111–115.
- [WEYL, H.] *The classical groups* (Princeton University Press, Princeton, NJ., 1946).

(Beérkezett: 1990. március 1.)

DOMOKOS GÁBOR
BME ÉPÍTÉSZMÉRNÖKI KAR SZILÁRDSÁGTANI TANSZÉK
1111 BUDAPEST, MŰEGYETEM RKP. 1-3.

INVESTIGATION OF THE ROTATIONAL SYMMETRIES OF PLANAR, CONSERVATIVE TENSOR FIELDS BY ELEMENTARY METHODS

G. DOMOKOS

Discrete rotational symmetries of polynomials in two variables are examined. One of the main propositions is that any polynomial can be constructed as a linear combination of C_n -symmetric kernels (depending only on n) and functions with continuous circular symmetry. The results are generalized for conservative tensor fields, the possible symmetry groups of homogeneous polynomials and constant tensor fields are identified. Based on these results the necessary degree of the *Taylor-approximation* of an operator is determined on the condition that the *truncated Taylor series* should share the symmetry group of the approximated operator. The applicability of the theorems is illustrated on examples from civil engineering mechanics, both analytical and numerical results are displayed.

OPTIMÁLIS RAGADOZÁSI STRATÉGIA A RAGADOZÓ-ZSÁKMÁNY VISZONYBAN

SZEMŐK ÁRPÁD

Szeged

Ismertetjük a *Volterra-Lotka ragadozó-zsákmány modellt*. Bevezetjük a ragadozási stratégia egy igen egyszerű definícióját, és egy több szempontból racionálisabb (a differenciálegyenletek szempontjából pedig általánosabb) modellt használunk arra, hogy ökológiai szempontból optimális ragadozási stratégiát keressünk. Grafikusan igen egyszerűen interpretálható optimalitási feltételeket kapunk. Eredményeink jó része az n ragadozó — 1 zsákmány esetre is átvihető.

1. A legegyszerűbb ragadozó-zsákmány modell

Az első világháború alatt, a háborús cselekmények következtében mintegy négy éven át erősen korlátozott volt a halászat az *Adriai tengeren*. A háború után készült mérések azt mutatták, hogy elszaporodtak a ragadozó halak a tengerben, és egyidejűleg lecsökkent bizonyos értékes növényevő halfajok száma. A statisztikai adatok elemzésére és a változások magyarázatára az olasz halbiológusok VITO VOLTERRÁT kérték fel. E munka eredményeképpen született meg a *Volterra-Lotka féle ragadozó-zsákmány modell* [3].

A zsákmányul szolgáló fajt mint növényevőt képzeljük el és feltételezzük, hogy számára korlátlan mértékben található táplálék. A ragadozó fajról feltesszük, hogy az adott és zártnak képzelt környezetben az előbbi faj képezi az egyetlen táplálékforrást.

Jelölje $x(t)$ illetve $y(t)$ a zsákmány ill. a ragadozó mennyiségét (egyedeinek számát) a t időpillanatban. A számuk helyett vehetnénk a „népsűrűségüket”, is ill. az össztömegüket is, a modellen ez nem változtatna. Feltételezzük, hogy a fajok létszáma olyan nagy, hogy rövid időtartamok alatti megváltozásuk folytonosnak tekinthető, más szóval feltételezzük, hogy $x(t)$ és $y(t)$ a t időnek folytonosan differenciálható függvénye.

Ha ragadozó nincs jelen, vagyis $y = 0$, akkor a legkézenfekvőbb feltevés az, hogy a zsákmányfaj természetes szaporodási rátája (számának időegység alatti megváltozása) állandó:

$$(1.1) \quad x'/x = \alpha \quad (\alpha > 0 \text{ állandó}).$$

A szerző megköszöni DR. HATVANI LÁSZLÓ kritikus megjegyzéseit, segítségét valamint DR. GALLÉ LÁSZLÓ biológus észrevételeit.

Hasonló módon, ha nincs zsákmányállat jelen, vagyis $x = 0$, akkor a legegyszerűbb feltevés az, hogy a ragadozó faj táplálék hiányában állandó mortalitási rátával kihal, vagyis $y'/y = -\gamma$, ahol $\gamma > 0$ állandó.

Ha a két faj együttesen van jelen az ökológiai környezetben, akkor a legegyszerűbb feltevés nyilván az, hogy az egyik faj természetes szaporodási rátája a másik faj számától függ, és ez a függés lineáris. E feltevések következménye a *Volterra-Lotka modell*:

$$\begin{aligned}x' &= x(\alpha - \beta y) \\ y' &= y(-\gamma + \delta x)\end{aligned}$$

Ezen egyenletrendszer viszonylag egyszerűen vizsgálható a pozitív síknegyedben. A rendszernek két egyensúlyi helyzete van: az $(x = 0, y = 0)$ ill. az $(x = \gamma/\delta, y = \alpha/\beta)$ pont. Bebizonyítható, hogy a többi megoldás periódikus a belső egyensúlyi helyzet körül — ugyanazzal a periódussal. A modell — bár nagyon leegyszerűsített — igen impozáns magyarázatot adott a lehalászásból és a lehalászás hiányából adódó arányeltolódásokra [3].

2. A Volterra-Lotka modell kritikája, általánosítása

A legkézenfekvőbb kifogás az, hogy $y = 0$ esetén a modell $x'/x = \alpha$ -ra redukálódik. Ezen differenciálegyenlet megoldása: $x(t) = x(0)e^{\alpha t}$. Erre a megoldásra alapozta THOMAS MALTHUS népesedési elméletét. Nem tűnik túlzottan realiztuskusnak az a modell, amely szerint a „magára hagyott faj” létszáma az időben exponenciálisan nő és a végtelenhez tart. Ez a hiba könnyen kiküszöbölhető. Tegyük fel, hogy a természetes szaporodási ráta a faj egyedszámának növekedésével csökken (esetleg stagnál), és ha a szám eléri azt az értéket, amennyit az ökológiai környezet tartósan fenntartani képes, akkor zérussá válik. (Ezt a számot jelöljük K -val: K — eltartóképesség.) Ekkor (1.1) helyett a következő összefüggést kapjuk: $x'/x = g(x)$, ahol $g(x)$ folytonosan differenciálható $g(0) > 0$, $g'(x) \leq 0$ és létezik $K > 0$: $g(K) = 0$, $g(x) > 0$, ha $x < K$ [3].

Ekkor a *Volterra-Lotka modell* így módosul:

$$\begin{aligned}x' &= x(g(x) - \beta y) \\ y' &= y(-\gamma + \delta x),\end{aligned}$$

ahol $g(x)$ az előbb leírt tulajdonságokkal bír. Ezt másképp is felírhatjuk:

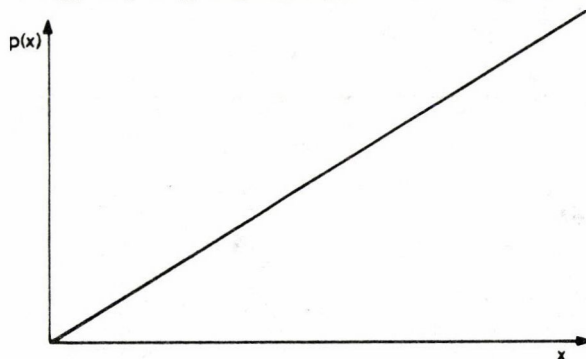
$$\begin{aligned}x' &= xg(x) - yp(x) \\ y' &= y(-\gamma + bp(x)),\end{aligned}$$

ahol $p(x) = \beta x$ és $b = \delta/\beta$.

Egy ragadozó- (vagy parazita-) populáció táplálkozási szokásainak leírására igen hasznos lehet a $p(x)$ „ragadozási” függvény [4], mely nem más mint egy ragadozó által egységnyi idő alatt elfogyasztott (megtámadott) zsákmányállatok átlagos mennyisége, amennyiben egy területegységen x mennyiségű zsákmányállat tartózkodik. Azaz:

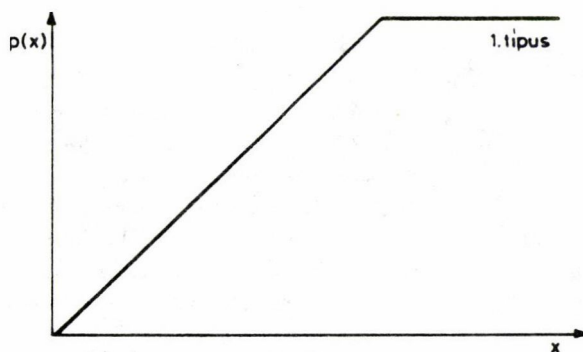
$$p(x) = \frac{\text{a ragadozók által elejtett zsákmányállatok mennyisége}}{(\text{a ragadozók száma}) \cdot (\text{eltelt idő})}$$

A ragadozási függvény nagyon jól szemléltethető a ragadozási görbével [4].



1. ábra

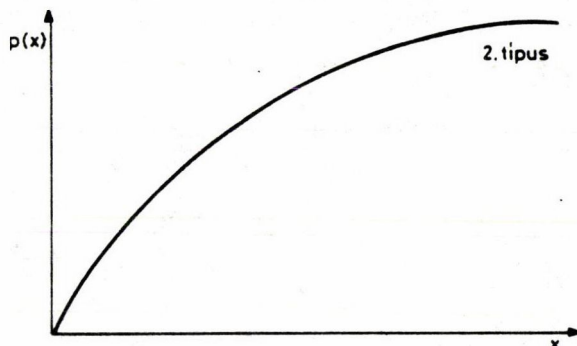
Az 1. ábrán a legegyszerűbb ragadozási görbe, a *Volterra–Lotka féle* található. Látszik, hogy ekkor az elejtett zsákmányállatok mennyisége egyenes arányban áll azok létszámával. A 2., 3. és 4. ábrán látható típusú ragadozási görbéket figyelték meg mindezekig az ökológusok [5,7].



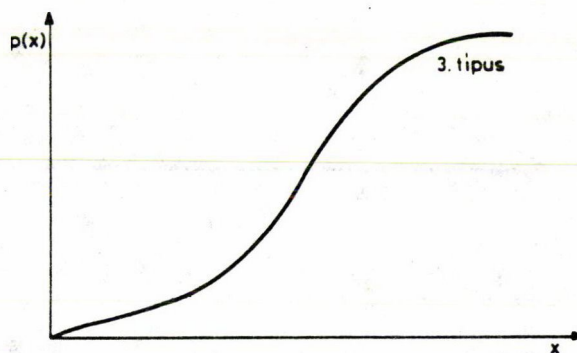
2. ábra

Az 1. pl. típusnál a ragadozás lineárisan nő egy határig. Ez tipikus táplálkozási mód pl. a ceteknél és minden olyan állatnál, mely vízből szűri a táplálékát. A 2. típusú görbe konkáv, monoton növekvő. Ez tipikus pl. a gerinctelen ragadozóknál és parazitáknál. A 3. típusú görbe első szakasza konvex, a második szakasza konkáv.

Ez nagyon gyakori a gerinces ragadozóknál, de ízeltlábú ragadozóknál és parazitáknál is megfigyelhető. Mindhárom típusú függvény korlátos.



3. ábra



4. ábra

Visszatérve a *Volterra-Lotka modell* kritikájára, rögtön szembetűnik, hogy a *Volterra-Lotka ragadozási-függvény* nem korlátos. A gyakorlatban elképzelhetetlen, hogy a ragadozófaj „étvágya” korlátlan legyen, és az sem szükségszerű, hogy $p(x)$ -nek akárcsak egy szakasza is lineáris legyen. Ezért, ha elhagyjuk a $p(x) = \beta x$ feltételt, akkor a modellnek egy ésszerű általánosítását kapjuk.

Mit célszerű $p(x)$ -ről feltenni [4]? Ha nincs zsákmány, akkor a ragadozó képtelen táplálékot szerezni: $p(0) = 0$. Az is teljesen reális feltevés, hogy a zsákmányállatok számának növekedésével az elejtett állatok száma is nő egy határig (az eltartóképessegig), de e határ felett sem csökken az elejtett állatok száma a zsákmány növekedésével.

Viszont a $p(x)$ korlátos, ami azt jelzi, hogy egy határ felett már a ragadozó saját fizikai képességei szabnak határt az elejtett állatok száma növekedésének.

Tehát az általunk használt modell [4]:

$$(2.1) \quad \begin{aligned} x' &= xg(x) - yp(x) \\ y' &= y(-\gamma + bp(x)) \end{aligned}$$

ahol $x, y \geq 0$ $g(0) > 0$ $g'(x) \leq 0$ $\exists K > 0 : g(K) = 0$ ($g(x) > 0$, ha $x < K$), $p(0) = 0$, p monoton nemcsökkenő, p folytonosan deriválható $[0, K)$ -n és $p'(x) > 0$ a $[0, K)$ -n, $\lim_{x \rightarrow \infty} p(x) + L$ ($< \infty$), γ, b állandók.

3. Definíciók, tételek

Tekintsük a következő n -változós differenciálegyenletrendszert:

$$(3.1) \quad \begin{aligned} x'_1 &= f_1(x_1, \dots, x_n) \\ &\vdots \\ x'_n &= f_n(x_1, \dots, x_n) \end{aligned}$$

ahol x_1, \dots, x_n ismeretlen függvények. Használjuk a kényelmesebb vektoros jelölésmódot: $\mathbf{z} := (x_1, \dots, x_n)^T$ $\mathbf{f} := (f_1, \dots, f_n)^T$. \mathbf{Z} vektor $\|\mathbf{Z}\|$ normája legyen $\|\mathbf{Z}\| = (x_1^2 + \dots + x_n^2)^{1/2}$. (3.1) vektoros alakja: $\mathbf{z}' = \mathbf{f}(\mathbf{z})$. Tegyük fel, hogy $\mathbf{z} = 0$ (azaz $x_1 = 0, \dots, x_n = 0$) egyensúlyi helyzete (3.1)-nek.

3.1. Definíció. A $\mathbf{z} = 0$ egyensúlyi helyzetet (*Ljapunov értelemben*) *stabilnak* nevezzük [1], ha bármely $\varepsilon (> 0)$ -hoz létezik olyan $\delta = \delta(\varepsilon)$, hogy minden \mathbf{z}_0 esetén, melyre $\|\mathbf{z}_0\| < \delta$ a (3.1) egyenlet $\varphi(0) = \mathbf{z}_0$ kezdeti feltétel melletti $\varphi(t)$ megoldása az egész $t > 0$ félegyenesen folytatható, és kielégíti a $\|\varphi(t)\| < \varepsilon$ egyenlőtlenséget minden $t > 0$ -ra.

3.2. Definíció. A $\mathbf{z} = 0$ egyensúlyi helyzet (*Ljapunov értelemben*) *aszimptotikusan stabil* [1], ha stabil, és létezik $\sigma (> 0)$ úgy, hogy $\|\varphi(0)\| < \sigma$ esetén $\lim_{t \rightarrow \infty} \varphi(t) = 0$. Azaz ha az origóhoz „elég” közelről indítjuk a megoldást, az az origóhoz fog tartani.

3.3. Definíció. Egy aszimptotikusan stabil egyensúlyi helyzet *attraktivitási tartományán* [8] (vonzási körzetén) mindazon pontok halmazát értjük, melyből indítva a megoldásokat, azok az egyensúlyi helyzethez tartanak.

Globális aszimptotikus stabilitásról [8] akkor szoktak beszélni, ha az egyensúlyi helyzet aszimptotikusan stabil és a vonzási körzet: \mathbf{R}^n . Mi most azt az egyensúlyi helyzetet nevezzük globálisan stabilnak, mely aszimptotikusan stabil és vonzási körzete a differenciálegyenletrendszer jobboldala értelmezési tartományának belseje. Ezt az indokolja, hogy modellünk a pozitív síknegyedben (a többdimenziós modellek a pozitív térrészben) van értelmezve, hiszen negatív mennyiségű élőlényről nincs értelme beszélni. A határokat azért nem vesszük figyelembe, mert 0-mennyiségű egyed már semmilyen körülmények között sem képes a szaporodásra.

A differenciálegyenlet(rendszer)ek megoldása ritkán adható meg explicit formulával, viszont egyensúlyi helyzeteikről és ezek stabilitásáról akkor is szerezhetünk információkat, ha megoldást találni teljesen reménytelen. Ebben hasznos segéd-eszköz *Ljapunov direkt módszere* ill. az ebben használt *Ljapunov-függvény*.

3.4. Definíció. *Ljapunov-függvénynek* [1] nevezzük azt a V skalárértékű függvényt, mely a $0 \in \mathbb{R}^n$ egy környezetében értelmezett, pozitív definit (azaz az origóban 0, mindenütt máshol pozitív értékű), és amelyre $\langle \text{grad } V(\mathbf{x}), f(\mathbf{x}) \rangle \leq 0$ teljesül, (ahol $\text{grad } V$ a V gradiensvektora, azaz $\text{grad } V = (\partial V / \partial x_1, \dots, \partial V / \partial x_n)$ valamint $\langle \alpha, \beta \rangle$ a belső szorzatot jelöli).

A *Ljapunov-függvényekkel* kapcsolatos tételek közül most egy olyat ismertetünk, melyre a dolgozatban is szükségünk lesz — nevezetesen a *Barbasin-Kraszovszkij tételnek* [8] egy számunkra szükséges következményét:

3.1. TÉTEL. Legyen $H = \{(\mathbf{x}_1, \dots, \mathbf{x}_n) : \mathbf{x}_i > -\omega_i \quad \omega_i > 0 \quad i = 1, \dots, n\}$ az origót tartalmazó nyitott térrész, és $f : H \rightarrow \mathbb{R}^n$. Legyen V *Ljapunov-függvény* a H értelmezési tartományon. Ha

- (a) Az $M = \{(\mathbf{x}_1, \dots, \mathbf{x}_n) : \langle \text{grad } V, f \rangle = 0\}$ halmaz nem tartalmaz egyetlen teljes trajektóriát sem az origón kívül, és
- (b) $\lim_{|\mathbf{x}| \rightarrow \infty} V(\mathbf{x}) = \infty$,

akkor az origó a (3.1) differenciálegyenletrendszernek globálisan aszimptotikusan stabil egyensúlyi pontja.

A stabilitásnak a hétköznapi szóhasználatban is vannak fokozatai. Az ökológusok is beszélnek igen stabil, stabil és gyengén stabil rendszerekről. A most következőkben a stabilitás egy lehetséges mérőszámát vezetjük be.

3.5. Definíció. A (3.1) egyenletrendszer aszimptotikusan stabil egyensúlyi helyzetének stabilitását jellemezze azon λ szám, mely azon μ -k infimuma, amelyek esetén az egyensúlyi helyzet vonzási körzetéből indított megoldások $\lim_{t \rightarrow \infty} \frac{\|\varphi(t)\|}{e^{\mu t}} = 0$ -nak tesznek eleget. (Magyarán azon μ számok alsó határát jelölje λ , amelyekkel egy $\|\varphi(t)\| < K e^{\mu t}$ becslés igaz minden 0-hoz tartó megoldás esetén.) λ -t *Ljapunov-mutatónak* [2] hívjuk.

Aszimptotikusan stabil egyensúlyi helyzetek esetén nyilvánvalóan $\lambda \leq 0$. Szemléletünk azt sugallja, hogy stabilabbnak mondjuk azt a rendszert, melynek *Ljapunov-mutatója* kisebb. A *Ljapunov-mutató* meghatározását könnyíti meg az alábbi tétel:

3.2. TÉTEL. Tegyük fel, hogy f differenciálható, $f(0) = 0$, és az
$$A = \left(\frac{\partial f_i}{\partial x_j} \mid x_1 = \dots = x_n = 0 \right)_{\substack{i=1, \dots, n \\ j=1, \dots, n}}$$
 mátrix sajátértékeinek valós részei negatívak. Ekkor az $\mathbf{x}' = f(\mathbf{x})$ egyenlet *Ljapunov-mutatója* az A mátrix sajátértékeinek maximumával egyenlő [6].

4. Optimálási feltételek

Jelen dolgozatunkban optimális ragadozó-stratégiát keresünk. Optimálisnak fogjuk tekinteni azt a ragadozó-stratégiát, mely maximális túlélési esélyt biztosít a ragadozó-populáció számára; a lehető legkedvezőbb módon viseli el mind az általa fogyasztott zsákmánypopuláció kisebb és nagyobb ingadozásait, mind az őt ért katasztrófákat.

Ez most matematikai szempontból azt jelenti, hogy a (2.1) differenciálegyenlet-rendszerben, melyet most modellként használunk, a $p(x)$ kivételével a paramétereket rögzítettnek tekintjük, és olyan $p(x)$ -et konstruálunk, mely a következőkben felsorolandó kritériumoknak eleget tesz. A kritériumok az optimális ragadozó stratégia egyes lehetséges megfogalmazásai.

1. *Legyen $p(x)$ olyan, hogy létezzen belső egyensúlyi helyzet!* Ennek az a biológiai jelentése, hogy létezik olyan együttélés, amelynek során mind a zsákmány- mind a ragadozópopuláció létszáma időben állandó és pozitív. A belső egyensúlyi helyzetet jelölje: (x^*, y^*) .

2. *Legyen $p(x)$ olyan, hogy y^* maximális értéket vegyen fel!* Ez biológiailag azt jelenti, hogy két egyensúlyi helyzet közül azt tekintjük előnyösebbnek, mely a ragadozóknak nagyobb létszámot biztosít, hiszen egy nagyobb katasztrófa esetén így több esélye van arra, hogy nem pusztul el a populáció minden tagja, mint egyébként.

3. *Legyen $p(x)$ olyan, hogy az egyensúlyi helyzetben a Ljapunov-mutató minimális legyen!* Ennek az a biológiai jelentése, hogy ha csak kicsit téríti ki valamilyen külső hatás a rendszert egyensúlyi helyzetéből, akkor a lehető leggyorsabban áll vissza az eredeti állapotába. (Ez nagy matematikai szigorral csak úgy igaz, hogy megfelelően sok idő elteltével tetszőlegesen közelre tér vissza.)

4. *Legyen $p(x)$ olyan, hogy a belső egyensúlyi helyzet globálisan aszimptotikusan stabilis legyen!* Ez úgy interpretálható, hogy ha egy egyensúlyban lévő rendszert olyan külső behatás ér, melynek következtében nem tűnnek el sem a ragadozók sem a zsákmányállatok, akkor az egyensúlyi helyzet „visszaáll”. (Az előző zárójeles megjegyzés itt is érvényes.) Nem kérdéses, hogy a 3. és 4. kritérium miért fontos, hiszen igen sok külső hatás érheti (éri) az ökoszisztémát: katasztrófák, járványok, környezet szennyeződés, vadászat. Nagyon fontos, hogy képes-e és ha igen, akkor milyen gyorsan képes visszanyerni a rendszer az egyensúlyi állapotot.

5. Egyensúlyi helyzet létezése, y^* maximalitása (1. és 2. kritérium)

Az (x^*, y^*) pont a (2.1) egyensúlyi helyzete, ha $x(t) \equiv x^*$, $y(t) \equiv y^*$ megoldása (2.1)-nek. Belső egyensúlyi helyzetről beszélünk, ha az egyensúlyi helyzet a pozitív síknegyed belsejében található. Azaz (x^*, y^*) belső egyensúlyi helyzet, ha $x^*g(x^*) = y^*p(x^*)$ és $-\gamma + bp(x^*) = 0$. Azaz $p(x^*) = \frac{\gamma}{b}$ és $y^* = \frac{x^*g(x^*)}{p(x^*)}$. A két egyenletet

egybevetve: $y^* = \frac{x^*g(x^*)}{\gamma}b$. Mivel $g(x), b, \gamma$ adott, így y^* akkor maximális, ha $x^*g(x^*)$ maximális. $xg(x)|_{x=0} = 0$, $xg(x)|_{x=K} = 0$ és $(xg(x))'|_{x=0} = g(0) + 0g'(0) = g(0) > 0$.

Ez utóbbiból és g folytonos differenciálhatóságából az is következik, hogy $xg(x) > 0$ $x = 0$ egy környezetében is igaz lesz. Azaz $xg(x)$ $x = 0$ -ban zérus, ezt követően monoton növe „egy darabig”, $x = K$ -ban ismét zérus, és ezután már nem lesz pozitív. Tehát $xg(x)$ maximuma létezik, és azt a $(0, K)$ intervallumon veszi fel. A folytonos differenciálhatóság miatt létezik az utolsó maximumhely. Ez legyen x^\wedge . Tegyük fel azt is, hogy $x^* = x^\wedge$, azaz x^* a lehető legnagyobb olyan értéket veszi fel, melyre létezik egyensúlyi helyzet és y^* maximális.

Ez lehetett volna egy újabb kritérium is (a „+1 kritérium”), de mivel jelentősége csekélyebb a másik négy kritériuménál, így nem soroltuk oda. (Mindig jelezni fogjuk a későbbiekben, hogy mi az a tény, ami a +1 kritériumból származik.)

A +1 kritérium jelentése az, hogy az egyensúlyi helyzetek közül azt részesítjük előnyben, melyre a zsákmánylétszám a lehető legnagyobb, hiszen egy esetleges — zsákmányokat érintő — katasztrófa esetén ekkor a legnagyobb az esély arra, hogy a zsákmánypopuláció, mely a ragadozók egyetlen tápláléka, nem hal ki.

Ha p -t úgy konstruáljuk, hogy $p x^\wedge$ -ben $\frac{\gamma}{b}$ -vel legyen egyenlő, akkor ez szükséges és elegendő feltétele lesz annak, hogy a (2.1) egyenlet belső egyensúlyi helyzete létezzen és y^* maximális legyen és x^* a lehetséges x^* -ok maximuma legyen. A tény a konstrukcióból nyilvánvaló.

Tehát $p(x^\wedge) = \frac{\gamma}{b}$ teljesíti az 1.2 és (+1) kritériumot, ahol x^\wedge az $xg(x)$ (utolsó) maximumhelye.

6. Maximális lokális stabilitás (3. kritérium)

Tudjuk, hogy x^* $xg(x)$ -nek lokális maximumhelye is, így $(xg(x))'|_{x=x^*} = 0$. A korábban leírtak szerint meg kell határozni a *Ljapunov-mutatót*, és meg kell néznünk, hogy az mikor minimális.

A variációs mátrix:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix}_{x=x^*, y=y^*} = \begin{pmatrix} (xg(x))'|_{x=x^*} - y^*p'(x^*) & -p(x^*) \\ y^*bp'(x^*) & -\gamma + bp(x^*) \end{pmatrix} = \\ = \begin{pmatrix} -y^*p'(x^*) & -p(x^*) \\ y^*bp'(x^*) & 0 \end{pmatrix}$$

A variációs mátrixban szereplő összes mennyiség adott, meghatározott, kivéve a $p'(x^*)$ -ot. $p'(x^*)$ mely értékeire lesz a *Ljapunov-mutató* a legkisebb?

A variációs mátrix sajátértéai a következő egyenlet gyökei (λ változó):

$$(\lambda + y^*p'(x^*))(\lambda - 0) + p(x^*)y^*bp'(x^*) = 0$$

Vezessünk be új változókat!

$$x := (\lambda + 2\gamma)/\gamma, \quad a := (y^* p'(x^*) - 4\gamma)/\gamma, \quad (\text{korábbi: } \gamma = bp(x^*))$$

Ekkor a karakterisztikus egyenlet új alakja (γ^2 -tel való osztás után):

$(x - 2)(x + 2 + a) + a + 4$ azaz $x^2 + ax - a = 0$ Hurwitz kritériuma szerint ezen egyenlet gyökei valósrészei csak akkor lehetnek egyszerre negatívak, ha a polinom együtthatói pozitívak. Azaz $a > 0$ és $-a > 0$. Ez lehetetlen. Viszont akkor nem lesz egyik gyök valós része sem pozitív, ha az együtthatók nemnegatívak: Azaz: $a \geq 0$ és $-a \geq 0$. (Ez a tény a Hurwitz-kritérium bizonyításában implicite benne van.) Ez csak $a = 0$ -ra teljesül. Visszatérve az eredeti változókra ez azt jelenti, hogy a gyökök valósrészeinek maximuma nem lehet kisebb -2γ -nál, -2γ pedig akkor és csak akkor lehet, ha $p'(x^*) = 4\gamma/y^* = \left(\frac{4\gamma^2}{bx^*g(x^*)}\right)$.

7. Globális aszimptotikus stabilitás (4. kritérium)

Vezessünk be új változókat: $w_1 := x - x^*$, $w_2 := y - y^*$. Azért van erre szükség, hogy az (x^*, y^*) belső egyensúlyi pont stabilitásának kérdéseit visszavezessük a $(0, 0)$ egyensúlyi pont vizsgálatára, hiszen definícióink és tételeink erre az esetre vannak megfogalmazva. Ekkor az új változókkal a (2.1) egyenletrendszer így fest:

$$\begin{aligned} w_1' &= (w_1 + x^*)g(w_1 + x^*) - (w_2 + y^*)p(w_1 + x^*) \\ (7.1) \quad w_2' &= (w_2 + y^*)(-\gamma + bp(w_1 + x^*)) \end{aligned}$$

Azt szeretnénk, hogy a $(w_1, w_2) = (0, 0)$ pont globálisan aszimptotikusan stabil legyen, miközben (7.1) értelmezési tartománya: a $H = \{(w_1, w_2) : w_1 > -x^*, w_2 > -y^*\}$ lezártja.

7.1 TÉTEL. Amennyiben $p(x)$ -re az is igaz, hogy:

$$(x - x^*) \left[\frac{xg(x)}{p(x)} - y^* \right] \leq 0,$$

akkor az origó globálisan aszimptotikusan stabilis.

Bizonyítás. A $V(w_1, w_2) = -y^* \ln \left(\frac{w_2 + y^*}{y^*} \right) + w_2 - bp(x^*)(P(w_1 + x^*) - P(x^*)) + bw_1$ Ljapunov-függvényt használjuk, ahol $P(x) = \int_1^x \frac{1}{p(s)} ds$. A 3.1. tétel értelmében azt kell belátni a H halmazon, hogy:

- $M = \{(w_1, w_2) : \langle \text{grad } V, (w_1', w_2') \rangle = 0\}$ nem tartalmaz az origón kívül egyetlen teljes trajektóriát sem.
- $\lim_{w \rightarrow \infty} (w) = \infty$

Mindenekelőtt azt bizonyítjuk be, hogy V valóban *Ljapunov-függvény*.
 $V(0, 0) = 0$, erről behelyettesítéssel könnyen meggyőződhetünk.
 Nézzük meg V parciális deriváltjait:

$$\begin{aligned}\frac{\partial V}{\partial w_1} &= -bp(x^*) \frac{\partial P(w_1 + x^*)}{\partial w_1} + b = b \left(1 - \frac{p(x^*)}{p(w_1 + x^*)} \right) = \\ &= \frac{b[p(w_1 + x^*) - p(x^*)]}{p(w_1 + x^*)} \\ \frac{\partial V}{\partial w_2} &= -y^* \frac{\partial}{\partial w_2} \left(\ln \left(\frac{w_2 + y^*}{y^*} \right) \right) + 1 = -y^* \frac{y^*}{w_2 + y^*} \cdot \frac{1}{y^*} + 1 = \\ &= \frac{-y^*}{w_2 + y^*} + 1 = \frac{w_2}{w_2 + y^*}\end{aligned}$$

p monotonitási tulajdonságai miatt igaz, hogy létezik az origónak egy olyan U környezete, hogy ha $(w_1, w_2) \in U$, akkor a parciális deriváltak a w_1 ill. w_2 változók pozitív ill. negatív értékre pozitívak ill. negatívak, és 0-ban zérók. A $H \setminus U$ tartományon pedig a parciális deriváltak a w_1 ill. w_2 változók pozitív ill. negatív értékre nemnegatívak ill. nempozitívak. Ez azt jelenti, hogy V -nek az origó abszolút (egyetlen) minimumhelye a H tartományon. Tehát V pozitív definit a H -n.

$$\begin{aligned}\langle \text{grad } V, (w_1, w_2) \rangle &= \frac{\partial V}{\partial w_1} w'_1 + \frac{\partial V}{\partial w_2} w'_2 = \\ &= \frac{b[p(w_1 + x^*) - p(x^*)]}{p(w_1 + x^*)} [(w_1 + x^*)g(w_1 + x^*) - (w_2 + y^*)p(w_1 + x^*)] + \\ &+ \frac{w_2}{w_2 + y^*} [(w_2 + y^*)(-\gamma + bp(w_1 + x^*))] = \\ &= \frac{b[p(w_1 + x^*) - p(x^*)](w_1 + x^*)g(w_1 + x^*)}{p(w_1 + x^*)} - b(w_2 + y^*)[p(w_1 + x^*) - p(x^*)] + \\ &+ w_2(-\gamma + bp(w_1 + x^*)) = \\ &= \frac{b[p(w_1 + x^*) - p(x^*)](w_1 + x^*)g(w_1 + x^*)}{p(w_1 + x^*)} - \\ &- bw_2p(w_1 + x^*) - by^*p(w_1 + x^*) + bp(x^*) + bp(x^*)(w_2 + y^*) - \gamma w_2 + \\ &+ bw_2p(w_1 + x^*) = \left(p(x^*) = \frac{\gamma}{b} \right) = \\ &= \frac{b[p(w_1 + x^*) - p(x^*)](w_1 + x^*)g(w_1 + x^*)}{p(w_1 + x^*)} - by^*p(w_1 + x^*) + \gamma y^* = \\ &= \frac{b[p(w_1 + x^*) - p(x^*)](w_1 + x^*)g(w_1 + x^*)}{p(w_1 + x^*)} + by^*[p(x^*) - p(w_1 + x^*)] = \\ &= b[p(x^*) - p(w_1 + x^*)] \left[\frac{(w_1 + x^*)g(w_1 + x^*)}{p(w_1 + x^*)} - y^* \right]\end{aligned}$$

A tétel feltétele miatt:

$$(x - x^*) \left[\frac{xg(x)}{p(x)} - y^* \right] = w_1 \left[\frac{(w_1 + x^*)g(w_1 + x^*)}{p(w_1 + x^*)} - y^* \right] \leq 0,$$

ezért nyilván igaz az, hogy $\langle \text{grad } V, (w'_1, w'_2) \rangle \leq 0$. Tehát V *Ljapunov-függvény*.

(a) bizonyítása: Az előző számolásokra rápillantva látszik, hogy létezik az origónak egy olyan kis U környezete, hogy $U \cap M = \{(w_1, w_2) : w_1 = 0, (w_1, w_2) \in U\}$, azaz az M halmaz az origó környékén a w_2 tengely egy darabja. Valóban, hiszen $w_1 \neq 0$ esetén $\langle \text{grad } V, (w'_1, w'_2) \rangle = 0$ akkor és csak akkor, ha $xg(x)/y^* = p(x)$. Ez azért lehetetlen, mert $x^* xg(x)/y^*$ -nek maximumhelye $p(x)$ -nek pedig növekedési pontja.

Az M nem tartalmazhat teljes trajektóriát az origón kívül, hiszen erre a trajektóriára az origó egy környezetében az lenne igaz, hogy $w_1 = 0$ és $w'_1 = 0$. A (7.1) egyenletbe beírva $w_1 = 0$ -t azt kapjuk, hogy $w'_2 = 0$. Eszerint a megoldás: $(w_1, w_2) = (0, c)$. Ez amiatt lehetetlen, mert az egyetlen egyensúlyi helyzet: $(w_1, w_2) = (0, 0)$. Ez azt jelenti, hogy M tényleg nem tartalmaz más trajektóriát csak az origót.

(b) bizonyítása: Tudjuk, hogy V parciális deriváltjai nemnegatívak a pozitív síknegyedben, ezért itt V w_1 és w_2 változójában is monoton nemcsökkenő, így elegendő belátni, hogy $\lim_{w_1 \rightarrow \infty} V(w_1, 0) = \infty$ és $\lim_{w_2 \rightarrow \infty} V(0, w_2) = \infty$.

A parciális deriváltak:

$$\lim_{w_1 \rightarrow \infty} \frac{\partial V}{\partial w_1} \Big|_{w_2=0} = \frac{b(L - \frac{\gamma}{b})}{L} > 0 \quad \lim_{w_2 \rightarrow \infty} \frac{\partial V}{\partial w_2} \Big|_{w_1=0} = 1$$

Mivel egyik határérték sem 0, ezért V egyik változójában sem konvergálhat véges számhoz. Tehát valóban igaz, hogy:

$$\lim_{w \rightarrow \infty} V(w) = \infty.$$

Összefoglalva: Amennyiben az is igaz, hogy:

$$(x - x^*) \left[\frac{xg(x)}{p(x)} - y^* \right] \leq 0,$$

akkor $p(x)$ eleget tesz a 4. kritériumnak.

Azt kaptuk, hogy ha $p(x)$ függvény olyan, hogy:

$$(7.2) \quad \begin{aligned} \text{I.} \quad & \frac{xg(x)}{p(x)} - y^* \geq 0, \quad \text{ha } x < x^* \\ \text{II.} \quad & \frac{xg(x)}{p(x)} - y^* \leq 0, \quad \text{ha } x > x^* \\ & p(x^*) = \frac{\gamma}{b} \text{ és } p'(x^*) = \frac{4\gamma^2}{bx^*g(x^*)}, \end{aligned}$$

akkor a ragadozó stratégia a korábban felvázolt 4. kritérium szerint optimális.

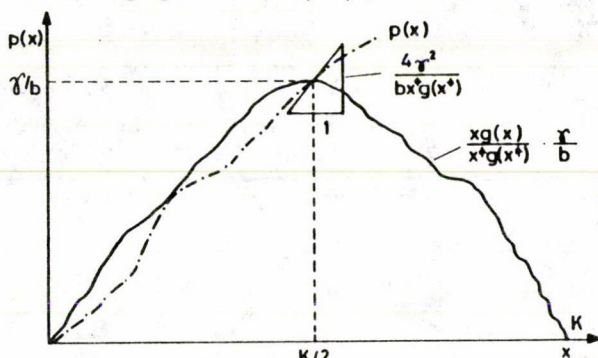
7.2. TÉTEL. A (7.2) feltételben a II. egyenlőtlenség elhagyható (a +1 kritérium miatt)

Bizonyítás. Tudjuk, hogy $\frac{x^*g(x^*)}{p(x^*)} - y^* = 0$. Azt is tudjuk, hogy $xg(x)$ utolsó maximumhelye: x^* (itt használjuk a +1 kritériumot), ezért $\frac{xg(x)}{p(x)} < \frac{x^*g(x^*)}{p(x^*)} (= y^*)$, ha $x > x^*$, ugyanis $xg(x) < x^*g(x^*)$ és $p(x) > p(x^*)$ ha $x > x^*$. Amivel az állítást beláttuk. Tehát az optimalitáshoz egy elegendő feltétel:

$$(7.3) \quad \text{I.} \quad \frac{xg(x)}{p(x)} - y^* \geq 0, \quad \text{ha } x < x^*$$

$$p(x^*) = \frac{\gamma}{b} \text{ és } p'(x^*) = \frac{4\gamma^2}{bx^*g(x^*)}.$$

Igen egyszerűen megfogalmazható (7.3) grafikus jelentése. (5. ábra)



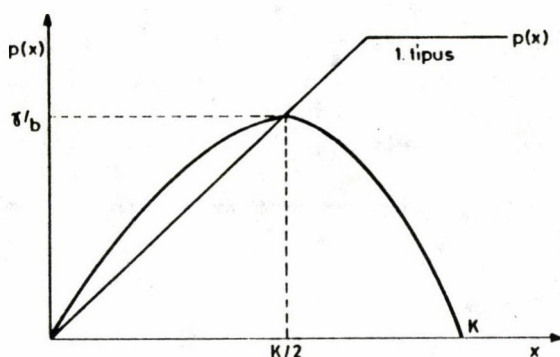
5. ábra

I.: $\frac{xg(x)}{p(x)} - y^* \geq 0 \iff \frac{xg(x)}{y^*} \geq p(x) \iff \frac{xg(x)}{x^*g(x^*)} \cdot \frac{\gamma}{b} \geq p(x)$. Azaz $p(x)$ a $[0, x^*]$ intervallumon az $\frac{xg(x)}{x^*g(x^*)} \cdot \frac{\gamma}{b}$ görbe fölé sosem megy. Az x^* helyen metszi a görbét, így teljesül ugyanis, hogy $p(x^*) = \frac{\gamma}{b}$; és p meredeksége az x^* helyen $\frac{4\gamma^2}{bx^*g(x^*)}$, ez jelenti ugyanis azt, hogy $p'(x^*) = \frac{4\gamma^2}{bx^*g(x^*)}$. Az x^* hely után a $p(x)$ görbe alakja érdektelen optimalitási feltételeink szempontjából. Ha a +1 kritérium nem lenne, akkor az utóbbi mondat úgy hangozna, hogy x^* után a $p(x)$ görbe sosem megy az $\frac{xg(x)}{x^*g(x^*)} \cdot \frac{\gamma}{b}$ görbe alá.

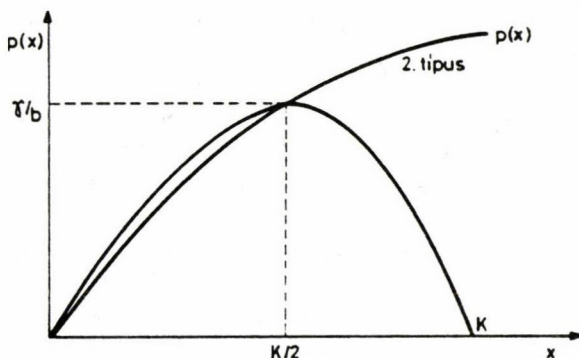
7.1. Megjegyzés. $g(x)$ gyanánt leggyakrabban az $r[1 - x/K]$ függvényt használják [7], melyben r a fajra jellemző konstans. K az élőhely eltartóképesége. Az $xg(x) = rx[1 - x/K]$ függvénynek egyetlen maximumhelye van: $\frac{K}{2}$. Ezért $x^* = \frac{K}{2}$.

Ez esetben: $\frac{xg(x)}{x^*g(x^*)} \cdot \frac{\gamma}{b} = \frac{r(1-\frac{x}{K})}{r\frac{K}{2}(1-\frac{1}{2})} \cdot \frac{\gamma}{b} = \frac{4x(1-\frac{x}{K})}{K} \cdot \frac{\gamma}{b}$ valamint $\frac{4\gamma^2}{bx^*g(x^*)} =$

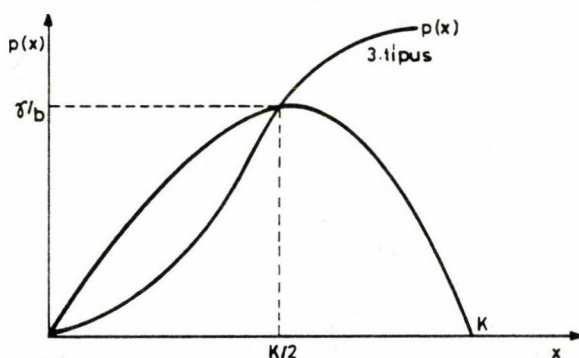
$$\frac{4\gamma^2}{rb\frac{K}{2} \cdot \frac{1}{2}} = \frac{16\gamma^2}{bKr}.$$



6a. ábra



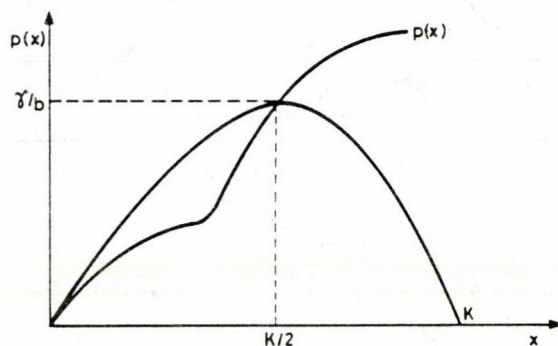
6b. ábra



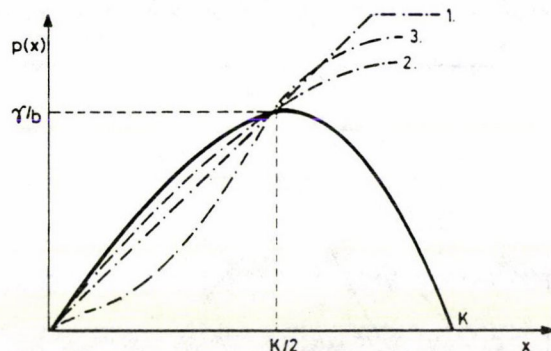
6c. ábra

Tehát, ha $g(x) = r[1 - x/K]$, akkor a kritériumok által leírt optimalitási feltételeknek megfelelő ragadozási stratégiát egy olyan $p(x)$ ragadozási függvény ír le, mely a $[0, \frac{K}{2}]$ intervallumon nem megy a $\frac{4x(1-\frac{x}{K})}{K} \cdot \frac{\gamma}{b}$ parabola fölé, és $\frac{K}{2}$ -ben $\frac{16\gamma^2}{bKr}$ meredekséggel metszi a parabolát. Ezek után már nem kétséges, hogy modellünkben — megfelelő paraméterekkel — mind a három típusú ragadozó magatartás lehet — az általunk felvázolt értelemben — optimális. (6. ábra)

De modellünk még a $g(x) = r[1 - x/K]$ esetben is lehetővé teszi, hogy egészen „vad” $p(x)$ ragadozási stratégiák létezzenek, melyek optimálisak a kritériumok szerint, mint például a 7. ábrán látható $p(x)$ függvény, mely először konkáv, majd konvex és ismét konkáv. Ez erős adhat azoknak az ökológusoknak, akik olyan mérési eredményekkel rendelkeznek, melyeket a hibaanalízis után sem képesek egyik típusba sem besorolni. De ez jelzés arra is, hogy érdemes újabb optimalitási feltételeken is gondolkodni.



7. ábra



8. ábra

7.2. Megjegyzés. Tegyük fel továbbra is, hogy $g(x) = r[1 - x/K]$. (Tudjuk, hogy ez az ökológusok számára nem nagy megszorítás.) Azt is tegyük fel, hogy $p(x)$ csak a felvázolt három típus valamelyikének reprezentánsa. Azt is tételezzük fel, hogy $p(x)$ a modellünkben a kritériumok szerint optimális. A 8. ábra is nyilvánvalóvá teszi, hogy $p(x)$ akkor és csak akkor lehet 1. típusú, ha a $\frac{K}{2}$ pontbeli meredeksége: $\frac{\gamma}{b} \frac{2}{K} p(x)$ akkor és csak akkor 2. típusú, ha $p'(\frac{K}{2}) < \frac{\gamma}{b} \frac{2}{K}$. $p(x)$ akkor és csak akkor 3. típusú, ha $p'(\frac{K}{2}) > \frac{\gamma}{b} \frac{2}{K}$. Mivel optimalitási kritériumaink értelmében $p'(\frac{K}{2}) = \frac{16\gamma^2}{bKr}$, ez azt jelenti, hogy $p(x)$ akkor és csak akkor 1. típusú, ha $\frac{16\gamma^2}{bKr} = \frac{\gamma}{b} \frac{2}{K}$ azaz ha $8\gamma = r$. Hasonlóan $p(x)$ akkor és csak akkor 2. típusú, ha $8\gamma < r$ és akkor és csak akkor 3. típusú, ha $8\gamma > r$.

Tudjuk, hogy γ a ragadozópopuláció mortalitási rátája, ami azt jelzi, hogy a táplálék nélkül maradt populáció milyen gyorsan pusztul el. r a ragadozóktól nem veszélyeztetett zsákmánypopuláció szaporodásának sebességéről ad információt. E két konstans aránya határozza meg (kizárólagosan) az optimális ragadozási stratégia típusát.

8. A modell általánosítása

A (2.1) modell egy lehetséges általánosítása [4] az a modell, melyben egy ragadozópopulációnak n darab zsákmánypopuláció jelenti (kizárólagos) táplálékforrását:

$$\begin{aligned}
 x'_1 &= x_1 g_1(x_1) - y p_1(x_1) \\
 x'_2 &= x_2 g_2(x_2) - y p_2(x_2) \\
 &\vdots \\
 x'_n &= x_n g_n(x_n) - y p_n(x_n) \\
 y' &= y[-\gamma + c_1 p_1(x_1) + c_2 p_2(x_2) + \dots + c_n p_n(x_n)]
 \end{aligned}
 \tag{8.1}$$

x_1, x_2, \dots, x_n a zsákmánypopulációk, y a ragadozópopuláció mennyiségét jelölik. $g_1(x_1), \dots, g_n(x_n)$ a zsákmánypopulációnak természetes szaporodási rátái, γ a ragadozópopuláció természetes mortalitási rátája.

$p_i(x_i)$ az i -edik zsákmánypopulációra vonatkozó ragadozási függvény ($p_i(x_i)$ definícióját úgy kapjuk, hogy $p(x)$ definíciójában a „zsákmány” szó helyébe „ i -edik zsákmány” szót helyettesítünk). A feltevések és azok indokai szó szerint átvihetők az általánosított esetre:

$x_1, \dots, x_n, y \geq 0$; $g_i(0) > 0$, g_i folytonosan deriválható, és létezik olyan $K_i > 0$, hogy $g_i(K_i) = 0$, ($g_i(x_i) > 0$, ha $x_i < K_i$), $p_i(0) = 0$, p_i monoton nemcsökkenő, p_i folytonosan deriválható $[0, K_i)$ -n és $p'_i(x_i) > 0$ a $[0, K_i)$ -n, $\lim_{x_i \rightarrow \infty} p_i(x_i) = L_i$ ($< \infty$) ($i = 1, \dots, n$) valamint γ, c_1, \dots, c_n konstansok.

Az $n = 1$ esetre megfogalmazott optimalitási feltételeknek olyan nyilvánvaló általánosítása kínálkozik, hogy a feltételek biológiai jelentését szükségtelen újra megfogalmazni. Most az optimális ragadozóstratégia-keresés azt jelenti, hogy a (8.1) differenciálegyenlet-rendszerben, melyet most modellként használunk, $p_i(x_i)$ ($i = 1, \dots, n$) kivételével a paramétereket (c, g_i, γ) rögzítettnek tekintjük, és olyan $p_i(x_i)$ függvényeket konstruálunk, mely a következőkben felsorolandó kritériumoknak eleget tesznek:

1. Létezzen belső egyensúlyi helyzet! (A belső egyensúlyi helyzetet jelölje:

$$(x_1^*, \dots, x_n^*, y^*)$$

2. y^* maximális értéket vegyen fel!
3. Az egyensúlyi helyzetben a *Ljapunov-mutató* minimális legyen.
4. A belső egyensúlyi helyzet globálisan aszimptotikusan stabil legyen.

8.a Egyensúlyi helyzet létezése, y^* maximalitása

Nyilvánvalóan $y^* = \frac{x_i^* g_i(x_i^*)}{p_i(x_i^*)}$ ($i = 1, \dots, n$) és $\gamma = \sum_{i=1}^n c_i p_i(x_i^*)$. A két egyenlőséget egybevetve: $\gamma = \sum_{i=1}^n c_i \frac{x_i^* g_i(x_i^*)}{y^*}$, amiből: $y^* = \sum_{i=1}^n \frac{c_i}{\gamma} x_i^* g_i(x_i^*)$. Mivel c_i, g_i, γ adott, y^* akkor maximális, ha minden $i = 1, \dots, n$ -re $x_i^* g_i(x_i^*)$ maximális. Már tudjuk, hogy $i = 1, \dots, n$ -re $x_i g_i(x_i)$ maximuma létezik, és a maximumhely a $(0, K_i)$ intervallumra esik. Ismét bevezetjük a +1 feltételt, ugyanis minden i -re létezik utolsó maximumhely.

Tehát x_i^* legyen $x_i g_i(x_i)$ utolsó maximumhelye, és $1/y^* := \sum_{i=1}^n \frac{c_i}{\gamma} x_i^* g_i(x_i^*)$.

$p_i(x_i^*) = \frac{x_i g_i(x_i^*)}{y^*}$ legyen. Ez esetben létezik belső egyensúlyi helyzet úgy, hogy y^* maximális és a +1 kritérium is teljesül.

8.b Globális aszimptotikus stabilitás

Vezessünk be új változókat: $w_i := x_i - x_i^*$, $u := y - y^*$. Ekkor (8.1) így fest:

$$w_1' = (w_1 + x_1^*)g_1(w_1 + x_1^*) - (u + y^*)p_1(w_1 + x_1^*)$$

$$w_2' = (w_2 + x_2^*)g_2(w_2 + x_2^*) - (u + y^*)p_2(w_2 + x_2^*)$$

$$\vdots$$

$$w_n' = (w_n + x_n^*)g_n(w_n + x_n^*) - (u + y^*)p_n(w_n + x_n^*)$$

$$u' = (u + y^*)[-\gamma + c_1 p_1(w_1 + x_1^*) + c_2 p_2(w_2 + x_2^*) + \dots + c_n p_n(w_n + x_n^*)]$$

8.1. TÉTEL. Ha

$$w_i \left[\frac{(w_i + x_i^*)g_i(w_i + x_i^*)}{p_i(w_i + x_i^*)} - y^* \right] \leq 0,$$

akkor az origó globálisan aszimptotikusan stabilis.

Bizonyítás. A bizonyítás az $n = 1$ esettel teljesen analóg. A $V(w_1, \dots, w_n) = -y^* \ln \left(\frac{u+y}{y^*} \right) + u + \sum_{i=1}^n c_i \{w_i - p_i(x_i^*)[P_i(w_i + x_i^*) - P_i(x_i^*)]\}$ Ljapunov-függvényt használjuk, ahol $P_i(x_i) = \int \frac{1}{p_i(x)} dx$.

8.1 Megjegyzés. A feltételek közül a Ljapunov-mutatóra vonatkozó feltétel az egyetlen, amelynek teljesüléséről nem beszélhetünk az $n = 1$ esethez analóg módon, hiszen a sajátértékek egy $n + 1$ -ed fokú polinom gyökeiként adódnak. A szerző sejtése az, hogy a Ljapunov-mutató minimális értéke -2γ , és ez csak akkor következik be, amikor $p_i'(x_i^*) = \frac{4\gamma}{y^*}$. (A sejtés alapja az, hogy az állítást $n = 2$ -re és bizonyos speciális esetekben $n > 3$ -ra is sikerült belátni.)

8.2. *Megjegyzés.* Mindkét modellre megfogalmaztunk egy-egy tételt, mely eleendő feltételt biztosít arra, hogy az origó globálisan aszimptotikusan stabilis legyen. (7.1., 7.2. tétel) Kérdés, ezek vajon szükséges feltételek-e.

8.3. *Megjegyzés.* A 7.2. megjegyzésben kaptunk egy összefüggést a ragadozási görbék típusa és a születési ill. mortalitási ráták aránya között. Igazolják-e ezt az ökológiai megfigyelések?

9. PÉLDÁK RAGADOZÁSI GÖRBÉKRE

Típus	Ragadozó	Zsákmány	Forrás
1.	<i>Artemia</i> (rák)	algák	HOLLING (1965)
1.	<i>Daphnia</i> (rák)	algák	HOLLING (1965)
2.	<i>Thais</i> (csiga)	kagylók	MURDOCH (1969)
2.	<i>Ezentrus canadiensis</i> (parazita rovar)	levéldarázs lárvá	GRIFFITHS (1969)
2.	<i>Harnodria azyridus</i>	<i>Aphis craccizora</i>	MOGI (1969)
2.	<i>Syrphus</i>	<i>Psyllid</i> lárvá	CLARK (1963)
3.	<i>Stentor</i> (egysejtű)	<i>Euglena</i>	RAPPORT (1974)
3.	<i>Coccinella septempunctata</i>	<i>Brericary brassicae</i>	COCK (1977)
3.	<i>Aphidius uzbekistanicus</i>	<i>Hylophroides humilis</i>	DRAUNFIED (1975)

Ezeket az adatokat és továbbiakat MURDOCH, OATEN [7] és HASSEL [5] könyveiben találhatunk.

IRODALOM

- [1] ARNOLD, *Közönséges differenciálegyenletek* (Műszaki Könyvkiadó, Budapest, 1987).
- [2] ДЕМИДОВИЧ, Лекции по математический теории устойчивости (Наука, Москва, 1967).
- [3] FARKAS MIKLÓS, "Stabilis együttélés és bifurkációk a populációdinamikában", *Alkalmazott Matematikai Lapok* 10 (1984).
- [4] FREEDMAN, *Deterministic Mathematical Models in Population Ecology* (Marcel Dekker, New York, Basel, 1980).
- [5] HASSEL, *The Dynamics of Arthropod Predation Prey Systems* (Princeton University Press Princeton N.J., 1978).
- [6] HATVANI L., "Nem-autonóm differenciálegyenletrendszerek", *Alkalmazott Matematikai Lapok* 5 (1979), 1-48.
- [7] MURDOCH, OATEN, "Predation and Population Stability", *Advances in Ecological Researches* 9 (1975), 2-132.
- [8] ROUCHE, HABETS, LALOY, *Stabilitáscelmélet. A Ljapunov-féle direkt módszer* (Műszaki Könyvkiadó, Budapest, 1984).

(Beérkezett: 1989. július 6.)

SZEMÖK ÁRPÁD
JÓZSEF ATILA TUDOMÁNYEGYETEM, BÓLYAI INTÉZET
6720 SZEGED, ARADI VÉRTANÚK TERE 1.

OPTIMAL PREDATOR FUNCTIONAL RESPONSE

Á. SZEMÖK

In this paper a generalized predator-prey differential equation model is studied. We are looking for a predator response function which is optimal from certain (biologically realistic) aspects. We show that the three types of the predator functional response which are known by the ecologists, are optimal in this sense. Most of our results can be used in the one predator – n preys model.

MEGJEGYZÉSEK A GILMORE—GOMORY MÓDSZER GYAKORLATI ALKALMAZÁSÁHOZ

LENGYEL IMRE

Békéscsaba

A dolgozatban az anyagleszabási feladatok megoldására kidolgozott *Gilmore–Gomory* módszer olyan kiterjesztésével foglalkozunk, amikor egy származtatott feladatot ismételtelen megoldva nyerjük az eredeti feladat egy elfogadható megoldását és az üzemi körülményeket paraméterek segítségével szimuláljuk. A módosítások hatását a szakirodalomból vett egy- és kétdimenziós mintafeladatokon szemléltetjük. A számítógépes futtatások tapasztalatait a dolgozatban leírtakon alapuló, több vállalatnál alkalmazott SZABOPT programcsomag segítségével ismertetjük.

1. Bevezetés

Anyagleszabási, darabolási problémának nevezzük azt az esetet, amikor meghatározott méretű és darabszámú alakzatot, idomot kell kiszabni úgy, hogy a felmerülő költség (speciális esetben a hulladék) minimális legyen. Az anyagleszabási probléma lineáris programozási feladatra (LP) való visszavezetése és *Gilmore–Gomory* módszerrel (GGM) történő megoldása közismert [5], [6–9], [11], [14], [15], [18]. Ennek ellenére a közelmúltban megjelent [3] dolgozatban a szerzők a szakirodalomban fellelt 34 gyakorlati esettanulmányt elemezve LP-re való visszavezetést csupán 15 esetben találtak és ebből a GGM felhasználását 4 alkalommal. A GGM csekély számú gyakorlati alkalmazása főleg annak tudható be (holott a mikroszámítógépek elterjedésével az alkalmazások száma is ugrásszerűen megnőtt), hogy lineáris modellek segítségével az anyagleszabás üzemi feltételei csak durva közelítéssel írhatók le. Így a módszer gyakorlati alkalmazása esetén az alapmodell bizonyos módosítására mindenképpen szükség van.

Dolgozatunkban a GGM olyan kiterjesztését mutatjuk be, amely egy származtatott feladatra történő visszavezetés segítségével az anyagleszabás üzemi feltételeinek paraméterekkel történő általánosításán alapszik. A paraméterek módosítása és a feladat ismételt megoldása lehetővé teszi a GGM korábbiaknál szélesebb körű alkalmazását. A darabolási problémák közül a paraméterek szerepét az egydimenziós esetben (rudak, csövek nyírása stb.) és az egydimenziós esetre visszavezethető kétdimenziós leszabásoknál (téglalapból–téglalapot típusú ortogonális kétütemű *guillotine-vágás* [9]) mutatjuk be. A dolgozat második fejezetében a fogalmak, jelölések ismertetésével együtt vázoljuk a GGM alapgondolatait. A harmadik fejezetben ismertetjük a származtatott feladatot és az üzemi körülmények paraméterekkel

történő újszerű modellezését. A negyedik fejezetben mutatjuk be a paraméterek alkalmazását szakirodalomból vett mintapéldákon, miáltal az eredmények is összevethetővé válnak, valamint a származtatásra közlünk futásjellemzőket. Az ötödik fejezetben térünk ki röviden a számítógépes alkalmazások gyakorlati tapasztalataira (a dolgozat lezárásakor 12 vállalatnál használták a leírtakon alapuló programcsomagot).

2. A Gilmore–Gomory módszer

A tanulmányban használt jelöléseket az *egydimenziós* lezabás segítségével vezetjük be. Legyenek az L_k , Q_k , C_k pozitív számok a k -adik alapanyag hossza, darabszáma és nyilvántartási ára (Q_k egész; $k = 1, 2, \dots, p$). Jelöljék a h_i , N_i pozitív számok a kiszabandó i -edik idom hosszát és igényelt darabszámát (N_i egész; $i = 1, 2, \dots, m$). Tegyük fel, hogy a szabásmintában, lezabási összeállításban az idomok elhelyezkedése, sorrendje tetszőleges és az összes lehetséges szabásminta száma: n . Ekkor a c_j költségű j -edik szabásmintát az $(a_{1j}, a_{2j}, \dots, a_{mj})^T$ vektorral, az ún. *szabás-vektorral* modellezzük, ahol a_{ij} nemnegatív egész és van olyan k ($1 \leq k \leq p$), hogy $\sum_{i=1}^m h_i a_{ij} \leq L_k$.

Az előzőekben megadott

$$(2.1) \quad \{L_k, C_k, Q_k, h_i, N_i \mid k = 1, 2, \dots, p; i = 1, 2, \dots, m\}$$

anyagleszabási problémához a következő egészértékű programozási feladat (ILP) rendelkezhető:

$$(2.2) \quad \begin{aligned} 0 &\leq x_j, \quad x_j = \text{egész} \quad (j = 1, 2, \dots, n) \\ \sum_{j=1}^n a_{ij} x_j &\geq N_i \quad (i = 1, 2, \dots, m) \\ \sum_{j=1}^n \partial_{kj} x_j &\leq Q_k \quad (k = 1, 2, \dots, p) \\ \sum_{j=1}^n c_j x_j &\rightarrow \min. \end{aligned}$$

ahol

$$\partial_{kj} = \begin{cases} 1, & \text{ha a } j\text{-edik szabásminta a } k\text{-adik alapanyagból készült,} \\ 0, & \text{egyébként.} \end{cases}$$

Valós üzemi problémánál a (2.2) ILP explicit felírása és a feladat megoldására hatékony eljárás kidolgozása a feladat nagy méretei miatt gyakorlatilag lehetetlen. Viszont a feladat *folytonos* változatára GILMORE és GOMORY adott egy hatékony

megoldási eljárást [7–8], amely kétfázisú módosított szimplexmódszeren és oszlopgeneráláson alapszik. A számításokhoz csak a bázist kell tárolnunk, a feladat mátrixára nincs szükség. A k -adik alapanyagot felhasználó $(a_1, a_2, \dots, a_m)^T$ szabás-vektor bázisba vonhatóságának feltétele a

$$(2.3) \quad \sum_{i=1}^m \Pi_i a_i + \Pi_{m+k} - c_k > 0$$

egyenlőtlenség fennállása, ahol $\Pi_1, \Pi_2, \dots, \Pi_m, \Pi_{m+1}, \dots, \Pi_{m+p}$ az aktuális bázishoz tartozó táblázat árnyékárai. A (2.3) egyenlőtlenség azt sugallja, hogy ebből az alapanyagból egy bázisba vonható vektort, amely szabás-vektor, az alábbi háti-zsákfeladat (KP) segítségével állítsunk elő:

$$(2.4) \quad \begin{aligned} 0 &\leq a_i, \quad a_i = \text{egész} \quad (i = 1, 2, \dots, m) \\ \sum_{i=1}^m h_i a_i &\leq L_k \\ \sum_{i=1}^m \Pi_i a_i &\rightarrow \max. \end{aligned}$$

Amennyiben a (2.4) feladatnak $k = 1, 2, \dots, p$ melletti megoldásai közül egy sem elégíti ki a (2.3) egyenlőtlenséget, akkor a szimplextáblázat a folytonos feladat optimális megoldását mutatja, amelyből a (2.2) ILP egy közelítő megoldását általában a pozitív multiplicitások egészértékűvé kerekítéséből nyerjük. A (2.2) ILP nem függ a felszabandó anyagok dimenzióitól, a leszabás sajátosságait, a műszaki és a szervezési feltételeket a szabás-vektor megkonstruálásakor kell figyelembe venni.

3. A GGM gyakorlati alkalmazásának szempontjai

A GGM gyakorlati alkalmazhatóságának legnagyobb akadálya az, hogy a (2.2) feladatra kapunk ugyan egy matematikai értelemben kielégítő megoldást, azonban az így nyert szabásminták ritkán felelnek meg a leszabás konkrét üzemi feltételeinek, azaz nem fogadhatók el [3], [12], [14]. A GGM elsősorban az anyagköltséget (illetve a keletkező hulladékot) minimalizálja, a szabásmintáktól nemlineárisan függő egyéb költségeket és szervezési elvárásokat (gyártásütemezés, a keletkező idomok készletezése, a selejtekt gyors pótlása, a szabászgépek egyenletes leterhelése stb.) az alapmodellben nem lehet figyelembe venni. Sőt olyan szubjektív, vagy egyéb nem formalizálható körülmények is döntőek lehetnek, amelyek a modellezés során nem ismertek.

Azért, hogy a GGM-t az üzemi körülmények minél szélesebb körénél alkalmazhassuk, a (2.4) KP-t több feltétellel kibővítjük és parametrizáljuk. A nem formalizálható üzemi körülményeket is a paraméterek vezérlésével szimuláljuk. A

(2.2) feladatot a paraméterek eltérő értékei mellett ismételtlen megoldjuk, ezáltal szabási tervvariánsok közül választható ki a konkrét anyagleszabási problémára egy, a gyakorlatban is elfogadható megoldás. A paraméterek vezérléséhez helyi üzemi tapasztalatokon alapuló *alkalmazói stratégiát* kell kidolgozni. A (2.2) feladat ismételt megoldását egy *származtatott feladatra* történő visszavezetéssel gyorsítjuk meg.

3.1 A származtatott feladat

A GGM lényegében kétfázisú módosított szimplexmódszer, első lépésben egy megengedett, a másodikban egy optimális bázismegoldást kell generálnunk. Az első fázis elkerülhető, ha az eredeti anyagleszabási problémából származtatunk egy olyan vele ekvivalens anyagleszabási problémát, amelyiknél az LP triviális bázisához megengedett megoldás tartozik, így az első fázist átugorva közvetlenül indítható a második fázis. Az eredeti feladat egy optimális megoldása a származtatott feladat optimális bázismegoldásából állítható elő [16].

Származtassunk a (2.1) anyagleszabási problémából valamelyik idom megkettőzésével, *részrendelések* létrehozásával egy új problémát. A továbbiakban az egyszerűség kedvéért tegyük fel, hogy az m -edik idomból, rendelésből csinálunk *kétféle* részrendelést. Ekkor a *származtatott probléma*:

$$(3.1) \quad \{L_k, C_k, Q_k, h'_i, N'_i \mid k = 1, 2, \dots, p; i = 1, 2, \dots, m+1\}$$

ahol $l'_i = l_i$ ($i = 1, 2, \dots, m$), $l'_{m+1} = l_m$ és $N'_i = N_i$ ($i = 1, 2, \dots, m-1$), $N'_m + N'_{m+1} = N_m$ ($N'_m > 0$, $N'_{m+1} > 0$ egészek).

Az eredeti és a származtatott problémához is egyértelműen rendelhető egy-egy LP, amelyek *ekvivalenciáján* azt értjük, hogy lehetséges megoldásaik között megfeleltetés létesíthető úgy, hogy a célfüggvényértékek megegyezzenek. A (2.2) feladatból felbontási szabállyal kapjuk meg a származtatott feladatot.

A *felbontási szabály* a következő: bővítsük a (2.2) feladatot úgy, hogy az m -edik idomból kettőt készítünk a darabszám kettéválasztásával. Ekkor az LP-ben az idomokhoz $m+1$ feltétel tartozik, amelyek jobb oldala:

$$N'_i = N_i \quad (i = 1, 2, \dots, m-1), \quad N'_m + N'_{m+1} = N_m \quad (N'_m > 0, \quad N'_{m+1} > 0 \text{ egészek})$$

Az $(a_{1j}, a_{2j}, \dots, a_{mj})^T$ j -edik szabás-vektorból képezzük az összes olyan lehetséges $m+1$ -dimenziós

$$(a'_{1js}, a'_{2js}, \dots, a'_{mjs}, a'_{m+1js})^T$$

szabás-vektort (jelölje az indexhalmazukat: S_j , $s \in S_j$), amelyre:

$$\begin{aligned} a'_{ijs} &= a_{ij} \quad (i = 1, 2, \dots, m-1) \\ a'_{mjs} + a'_{m+1js} &= a_{mj} \quad (a'_{mjs} \geq 0, \quad a'_{m+1js} \geq 0 \text{ egészek}). \end{aligned}$$

A (3.1) származtatott anyagleszabási probléma LP-je a (2.1) feladatból a felbontási szabály alkalmazásával a következő:

$$\begin{aligned}
 & 0 \leq y_{js} \quad (j = 1, 2, \dots, n; s \in S_j) \\
 & \sum_{j=1}^n \left(\sum_{s \in S_j} a'_{ij,s} y_{js} \right) \geq N'_i \quad (i = 1, 2, \dots, m+1) \\
 (3.2) \quad & \sum_{j=1}^n \partial_{kj} \left(\sum_{s \in S_j} y_{js} \right) \leq Q_k \quad (k = 1, 2, \dots, p) \\
 & \sum_{j=1}^n c_j \left(\sum_{s \in S_j} y_{js} \right) \rightarrow \min
 \end{aligned}$$

A (2.2) és a (3.2) feladatok ekvivalenciája a felbontási szabályt alkalmazva egyszerűen belátható. Ennek alapján bebizonyítható az is, hogy bármelyik idom-feltétel véges számú felbontása indukciót alkalmazva elvégezhető úgy, hogy a (2.2) feladattal ekvivalens származtatott feladatokat nyerünk.

A fenti visszavezetéssel az a célunk, ha a homogén szabás-vektorok (homogén egy szabás-vektor, amelynek csak egy eleme pozitív) nem alkotják a (2.2) feladat egy lehetséges bázismegoldását, akkor minél kevesebb részrendelésekre bontással előállítsunk egy olyan származtatott feladatot, amelynek homogén szabás-vektorokat tartalmaz (azaz triviális) egy induló bázismegoldása.

Más lineáris programozási modelleknél a fenti származtatást csak akkor érdemes alkalmazni, ha az LP mátrixának ismeretére explicite nincs szükség. Például oszlopgenerálás esetén, amikor a bázisba bevonható vektorokat valamilyen segédfeladat segítségével elő tudjuk állítani [1], egyébként a származtatott feladat mérete a (2.2) feladathoz képest jelentősen megnőhet. Oszlopgeneráláskor azonban csak a bázisra, illetve a szimplextáblára van szükség, a (3.2) feladat mátrixának tárolására nem, ezért a számítógépes memóriaigény alig növekszik (a futtatások jellemzőire a 4. fejezetben térünk ki).

3.2 Az általánosított hátizsákfeladat paraméterei

A felhasznált paraméterek egyrészt a *műszaki* adottságokat modellezik (pl. a vágófejek maximális és minimális számát, ha ezek a fejek egyszerre vágnak), másrészt az *üzemi körülmények szimulálására* adnak lehetőséget. A (2.4) KP kiterjesztésénél a feltételeket felülről korlátozó és egyféle alapanyagot figyelembe vevő szakirodalmi eredményeket ([4], [10], [13], [24]) általánosítottuk úgy, hogy a feltételeket alulról is korlátozzuk és többféle alapanyagot veszünk figyelembe. A [19] dolgozat eredményeit kiterjesztjük kétdimenziós esetre és röviden foglalkozunk az oszlopgenerálás néhány gyakorlati kérdésével.

Kiterjesztésünknel az általánosított KP *egydimenziós* esetben a k -adik alapanyagra a következő nemlineáris többfeltételes feladat:

$$(3.3) \quad l_{ik} \leq a_i \leq u_{ik}, \quad a_i = \text{egész} \quad (i = 1, 2, \dots, m)$$

$$(3.4) \quad \sum_{i=1}^m h_i a_i \leq L_k$$

$$(3.5) \quad R_{lk} \leq \sum_{i=1}^m a_i \leq R_{uk}$$

$$(3.6) \quad T_{lk} \leq \sum_{i=1}^m \text{signum } a_i \leq T_{uk}$$

$$(3.7) \quad \sum_{i=1}^m \Pi_i a_i \rightarrow \max.$$

ahol l_{ik} , u_{ik} ($i = 1, 2, \dots, m$) és R_{lk} , R_{uk} , T_{lk} és T_{uk} nemnegatív egészek (természetesen $R_{uk} \geq T_{uk}$ és $R_{uk} \geq u_{ik}$).

A (3.3) a változók alsó és felső korlátait tartalmazza, megfelelő egyedi korlátokkal a szabásminta-féleségek (ezáltal a vágófejek átállításának) száma is csökkenthető. A (3.5) feltétel a vágófejek minimális és maximális számára vonatkozik, az R_{lk} és R_{uk} paraméterekkel irányítható a szabásmintákban a darabszám egyenletessége, így a munkások leterheltsége is. A (3.6) feltétellel a szabásminták *bonyolultsága* vezérelhető, ezáltal a vágóasztalról történő leszedés, tárolás körülményeit is szimulálhatjuk. Valamint elkerülhető az, hogy sok kisméretű idomot tartalmazó, munkaigényes szabásminta váltakozzon csak nagyméretű idomokat tartalmazó, egyszerűen leszabható szabásmintával [20].

Kétütemű kétdimenziós guillotine-vágásnál először az alaplemezt valamelyik oldalára merőleges vágásokkal csikokra szabjuk fel. Ezeket a csikokat, mint egydimenziós anyagokat daraboljuk tovább téglalapokra, amelyek vagy a kért méretű idomok, vagy már csak egyféle idomokat tartalmaznak (*free two-stage cutting*). A bázisba bevonandó szabás-vektort is két lépésben adjuk meg. Először mindegyik csík-szélességhez meghatározzuk a legjobb összeállítást, majd ezekből a csikokból elkészítjük a szabásmintát (mindkét lépés során egydimenziós feladatokat kell megoldanunk).

Legyenek az L_k , W_k , Q_k , C_k pozitív számok a k -adik alaplemez hossza, szélessége, darabszáma és nyilvántartási ára (Q_k egész; $k = 1, 2, \dots, p$). Jelöljük a h_i , w_i , N_i pozitív számok a kiszabandó i -edik idom hosszát, szélességét és igényelt darabszámát (N_i egész; $i = 1, 2, \dots, m$). Tegyük fel, hogy $w_1 \leq w_2 \leq \dots w_m$. Ekkor a k -adik alapanyagból a bázisba bevonható szabás-vektort a következő két lépésben határozzuk meg, ahol a jelölések megegyeznek a (3.3–7) feladat jelöléseivel ([]- az egészrész jele):

1. lépés: elkészítjük a w_j szélességű legjobb csikot az alábbi feladat segítségével

vel:

$$\begin{aligned}
 l_{ik} &\leq b_{ij} \leq u_{ik}, \quad b_{ij} = \text{egész} \quad (i = 1, 2, \dots, j) \\
 \sum_{i=1}^j h_i b_{ij} &\leq L_k \\
 R_{lk} &\leq \sum_{i=1}^j b_{ij} \leq R_{uk} \\
 T_{lk} &\leq \sum_{i=1}^j \text{signum } b_{ij} \leq T_{uk} \\
 \sum_{i=1}^j \Pi_i \left[\frac{w_j}{w_i} \right]^- b_{ij} &\rightarrow \max.
 \end{aligned}
 \tag{3.8}$$

Legyen a célfüggvény maximuma: Π_j^* . A (3.8) feladatot $j = 1, 2, \dots, m$ -re meg kell oldani, az optimális megoldások legyenek: $b_{1j}^*, b_{2j}^*, \dots, b_{jj}^*$.

2. lépés: az 1. lépésben meghatározott csík-összeállításokból megkonstruáljuk a legjobb szabásmintát:

$$\begin{aligned}
 0 &\leq d_j, \quad d_j = \text{egész} \quad (j = 1, 2, \dots, m) \\
 \sum_{j=1}^m w_j d_j &\leq W_k \\
 R_{lk}^* &\leq \sum_{j=1}^m d_j \leq R_{uk}^* \\
 T_{lk}^* &\leq \sum_{j=1}^m \text{signum } d_j \leq T_{uk}^* \\
 \sum_{j=1}^m \Pi_j^* d_j &\rightarrow \max,
 \end{aligned}
 \tag{3.9}$$

ahol az $R_{lk}^*, R_{uk}^*, T_{lk}^*$ és T_{uk}^* pozitív egészek a (3.3–7) feladat paramétereinek megfelelő általánosításai csíkokra. Legyen $d_1^*, d_2^*, \dots, d_m^*$ a (3.9) feladat optimális megoldása. Ekkor a legelőnyösebb szabás-vektor komponensei:

$$a_i = \sum_{j=1}^m \left[\frac{w_j}{w_i} \right]^- b_{ij}^* d_j^* \quad (i = 1, 2, \dots, m)$$

Egy- és kétdimenziós esetben is a szabás-vektor megkonstruálásához a nemli-neáris többfeltételes feladatokat minden alapanyag-méretre meg kell oldani (kétdimenziós esetben ha az alapanyag elforgatható, akkor a feladatok száma megduplázódik). Többféle alapanyagnál kérdésként vetődik fel, hogy melyik k melletti

megoldást vigyük be a bázisba? A GGM alkalmazásainál általában a különbséget maximalizáló esetet veszik alapul:

$$(3.10) \quad \max_k \left\{ \sum_{i=1}^m \Pi_i a_i + \Pi_{m+k} - C_k \mid k = 1, 2, \dots, p \right\}$$

Az összehasonlító számítógépes futtatások során célszerűbbnek tűnt a legjobb fajlagos kihasználású alapanyaghoz rendelt szabás-vektor bázisba vitele:

$$(3.11) \quad \max_k \left\{ \left(\sum_{i=1}^m \Pi_i a_i + \Pi_{m+k} \right) / C_k \mid k = 1, 2, \dots, p \right\}$$

A futásidő szempontjából általában előnyösebb, ha nem keressük meg az általánosított KP-k optimális megoldását, hanem elfogadunk egy olyan vektort, amely a (2.3) egyenlőtlenséget kielégíti. A (3.3-7) feladatcsaládot megoldó leszámoló algoritmusból kilépünk, ha találunk egy olyan $a = (a_1, a_2, \dots, a_m)^T$ vektort, amelyre

$$(3.12) \quad \sum_{i=1}^m \Pi_i a_i + \Pi_{m+k} > e C_k$$

ahol $e > 1$ alkalmasan választott konstans. A (3.8) feladatnál a kiugrási feltétel:

$$(3.13) \quad \sum_{i=1}^j \Pi_i \left[\frac{w_j}{w_i} \right]^- b_{ij} > e C_k \left(\frac{w_j}{w_k} \right)$$

míg a (3.9) feladatnál ismét a (3.12) képletet alkalmaztuk.

4. Mintapéldák a paraméterek szerepére és futtatási eredmények a származtatásra

Előbb a GGM vezérlését két konkrét szakirodalmi példa alapján mutatjuk be, így módosításaink összehasonlíthatók az eredeti módszerrel, majd a származtatott feladatra közlünk futtatási eredményeket. A programokat IBM PC/XT számítógépen futtattuk le QUICKBASIC programozási nyelven.

Egydimenziós esetben *Gilmore és Gomory* egyik példájára ([8], 873. oldal, 2. táblázat) ismertetjük a futási eredményeket. Ott a szerzők egy alapanyagot adtak meg (a hossza 218 inch), mi pedig ötöt (az adatokat az 1. és 2. táblázat tartalmazza). A paraméterek eltérő értékein alapuló számítógépes futások jellemző adatai a 3. táblázatban találhatók. Az alapanyagokat először az A áron vettük figyelembe és a felhasznált alapanyag költségének minimalizálása volt a cél. Jelen esetben a kétfázisú módosított szimplexmódszerrel dolgoztunk, a származtatást az összehasonlíthatóság miatt nem vettük igénybe.

1. táblázat. Az egydimenziós példa rendeléseinek adatai

	hossz	darabszám	egyedi korlát u_{ik}
1.	81,00	4415	4
2.	70,00	291	2
3.	68,25	4765	5
4.	67,50	4827	4
5.	66,75	90	1
6.	66,00	691	3
7.	64,00	263	2
8.	63,75	141	2
9.	63,00	133	2
10.	60,00	390	2
11.	56,25	459	2
12.	56,00	343	2
13.	52,50	766	3
14.	52,00	58	1
15.	51,75	2736	4
16.	52,00	212	2
17.	50,00	720	3
18.	49,50	133	2
19.	46,50	529	3
20.	45,50	185	2
21.	44,50	94	2
22.	41,25	393	2
23.	38,50	47	1
24.	38,00	95	1
25.	35,00	411	2
26.	33,50	36	1
27.	33,00	273	2
28.	32,00	56	1
29.	31,50	171	2
30.	21,50	140	2

2. táblázat. Az egydimenziós példa alapanyagainak adatai

	hossz	darabszám	nyilvántartási ár	
			A.	B.
1.	267	10,000	267	267
2.	243	1,000	243	223
3.	218	1,000	218	195
4.	203	10,000	203	203
5.	185	10,000	185	185

3. táblázat. Futtatási eredmények az egydimenziós mintapéldára

	bázistranszformációk	hulladék /%/		futásidő /sec./	
		folyt.	egészért.	összes	KP
/a/	130	0,08	0,33	972	622
/b/	214	0,09	0,32	1.796	1.085
/c/	121	7,51	7,70	484	219
/d/	74	5,18	5,23	404	283
/e/	89	9,08	9,30	349	176
/f/	145	0,10	0,31	795	434

Az (a) esetben a vezérléstől eltekintettünk, minden feltételelt lazának adtunk meg ($l_{ik} = 0$, $R_{ik} = T_{ik} = 1$ és $u_{ik} = R_{uk} = T_{uk} = 30$), így ez az eset felel meg az eredeti GGM-nek. A további esetekben egy-egy paramétert erősen korlátozónak vettünk. A (b) esetben az egyedi felső korlátok szigorúak, az 1. táblázatban található az értékek. A (c) esetben az idomféleségek száma ($T_{uk} = 3$), míg a (d) esetben az idomok összesített darabszámát ($R_{uk} = 4$) korlátoztuk felülről. Az (e) esetben a (b), (c) és (d) esetek felső korlátait *együtt* vettük figyelembe. Az (f) esetet az (a) eset azon változata, amikor a (3.12)-ben szereplő konstans értéke $e = 1,5$.

A 3. táblázatból látható, hogy az egyedi felső korlátok alacsony értéke a hulladékot nem befolyásolja, csupán a futásidő nőtt meg közel a kétszeresére. Az optimális szabásmintáknak a száma az (a) esetben 30, míg a (b) esetben 27, azaz egyedi korlátokkal irányítható a szabásgépek átállítása. Az idomféleségek számának (c), valamint a szabásmintában szereplő idomok összesített darabszámának felülről történő korlátozása (d) a hulladékot jelentősen növelte, a futásidőt pedig csökkentette. Hasonló tapasztalható a felülről erősen korlátozott (e) esetben is, ami a lehetséges megoldások halmazának szűkülésére vezethető vissza. Természetesen R_{uk} és T_{uk} értékét növelve a hulladék csökken. Az iterációk számának növekedése ellenére az (a)-hoz képest a futásidő közel 20 %-kal kisebb az (f) esetben, amikor elfogadtuk a KP-nak egy közelítő megoldását. A bázistranszformációk száma megnőtt, azonban ez kevésbé növeli az összesített futásidőt, mivel a KP-k megoldása jóval időigényesebb a szimplextáblázat átszámításánál.

A 3. táblázat adatai a GGM (3.11) szerinti, a fajlagos értéket maximalizáló futtatásából származnak. Mind a hat esetben (3.10)-nek megfelelően is lefuttattuk a programokat. A futásidő a (3.11) képlet alkalmazásakor négyszer volt kedvezőbb, míg kétszer közel hasonló. Amikor az alapanyagokat a 2. táblázat B áráival vettük figyelembe, tehát a kedvező árú anyagok raktárkészlete szűkös volt, akkor a futásidő 30-40 %-kal megnőtt (természetesen az alacsonyabb fajlagos költségű alapanyagok mind fel lettek használva).

A (2.2) ILP egy közelítő megoldását a megfelelő folytonos LP optimális megoldásának egészértékűvé kerekítéséből nyerjük, ez a kerekítés a hulladék kb. 0,2%-os növekedésével jár együtt. Az LP célfüggvényértékének alakulását vizsgálva kiderül,

hogy a bázistranszformációk bizonyos száma után a megfelelő egészértékű megoldás alig javul, sőt néha romlik is. Például az (a) esetben már 88. transzformáció után alig rosszabb, sőt a 128. lépésnél még jobb is az egészértékű megoldás az LP alapján elfogadottnál. Ezért a transzformációk során célszerű tárolni az addigi legjobb egészértékű megoldást és szabás-vektorait. A célfüggvény a 88. transzformáció után alig csökkent (a hulladék 0,42 és 0,31 % között mozgott), amiből az is következik, hogy sok, a gyakorlatban azonos értékűnek tekinthető szabásminta-sorozat van, amelyek közül a formálisan nem modellezhető egyéb üzemi körülményeket is mérlegelve lehet egy elfogadhatót kiválasztani.

Az alsó egyedi korlátoknak (l_k) pozitív értéket választva a futásidő kismértékben megnőtt. Viszont csak egy-két idom alsó korlátjának célszerű pozitív értéket adni, mivel a lehetséges megoldások halmaza könnyen üressé válhat. Az alsó korlátokkal érhetjük el valamely idom, vagy idomok kidarabolását egy meghatározott alapanyagból.

Az üzemi gyakorlatban sűrűn felvetődik az *idomok kielégítésének folyamatossága* (lásd [17], [23]). Ez azt jelenti, ha egy idom szerepel egy szabásmintában, akkor a gyártásütemezés szerint soron következőkben is szerepeljen addig, amíg az igényelt darabszámban el nem készült. Matematikai értelemben ez azt jelenti, hogy az optimális megoldás pozitív multipllicitásaihoz tartozó szabás-vektorokból álló mátrix oszlopait (illetve sorait) át kell úgy rendezni, hogy a mátrix soraiiban a pozitív elemek szomszédosak legyenek. Ez a rendezés csak speciális feltételeket kielégítő mátrix esetében végezhető el [22], amely feltételeknek a darabolási feladat megoldásából nyert mátrix nem felel meg, ezért csupán a folyamatosság megsértésének minimalizálása jöhet szóba. MADSEN a szabás-vektorokból képzett szimmetrikus mátrix sávszélességének minimalizálására vezette vissza ezt a problémát (*bandwidth problem*) [21]. Az idomok kielégítésének folyamatosságával együtt az alapanyagok bekészítésének folyamatossága is felvetődik, az egymás után készülő szabásképek lehetőleg sokáig készüljenek ugyanazon anyagból. A két folyamatossági cél ellentmond egymásnak, ez egy nehéz NP-teljes feladat [23], amelynek gyakorlati megoldásával foglalkozó publikáció a dolgozat lezárásáig általunk nem ismert.

A GGM bemutatott kiterjesztésénél a folyamatosság megsértését a (3.5) és (3.6) feltételekben levő R_{uk} és T_{uk} paramétereknek alacsony értéket adva vezérelhetjük. Ezáltal a szabás-vektorokban kevés pozitív elem lesz, „ritka” mátrixot kapunk és így a rendezés könnyebben elvégezhető. Pl. az (a) esetben, amely az eredeti GGM többféle alapanyagra történő alkalmazásának felel meg, az idomok leszabása 9-szer, míg a (c) esetben 6-szor szakítódik meg. Mindkét esetben egy heurisztikus rendező eljárást alkalmaztunk.

Kétdimenziós esetben WANG egyik példáját ([24], 585. oldal, 4. táblázat) oldottuk meg kétütemű leszabásra. A számítógépes futások tapasztalatai, a vezérlések futásidőre és hulladékra gyakorolt hatása hasonló az egydimenziós esethez, ezért csak röviden vázoljuk a specifikus eltéréseket. Kétdimenziós esetben a futásidő 90–95 %-át a (3.8) és (3.9) feladatok megoldása teszi ki. Ezért a futásidő lényegesen meggyorsítható a (3.12) és (3.13) képletek alkalmazásával, amikor az első és a második lépés nemlineáris feladatainak is elfogadjuk valamely közelítő megoldását. A

felülről korlátozó paramétereket elég nagynak adva ennél a kétdimenziós példánál a hulladék 4,96 % (a bázistranszformációk száma: 33, a futásidő 814 sec.) míg a feltételeket megszigorítva 5,83 % lett (a bázistranszformációk száma 39, a futásidő: 1,143 sec.).

4.1 Futtatási eredmények a származtatásra

Ugyanazon anyagleszabási problémákat oldottuk meg az eredeti kétfázisú, valamint a származtatott feladat megoldásánál használt egyfázisú módosított szimplexmódszerrel. Egydimenziós leszabást vettünk alapul és két feladatcsoportot futtattunk le, az egyiknél $p = 5$ és $m = 15$, a másiknál $p = 10$ és $m = 30$. Mindkét csoportnál 10 feladatot oldottunk meg (természetesen az eredetit és a belőle származtatott feladatot is). Az alapanyagok és az idomok adatait a számítógép egyenletes eloszlású véletlenszám-generátorával állítottuk elő, ahol:

$$1400 \leq L_j \leq 3400 ; \quad 100 \leq Q_j \leq 1000 \\ 10 \leq l_i \leq 460 ; \quad 200 \leq N_i \leq 1200$$

Az alapanyagokból 2–4 volt olyan, amelyek készletét alacsonyra vettük és ezeknél az egységárat is csökkentettük, hogy nagyobb valószínűséggel szerepelhessenek az optimális szabásminták anyagaként. A többi alapanyag fajlagos árát azonosnak adtuk meg. A feladatcsoportok futási eredményeit a 4. táblázat tartalmazza. A hátizsákfeladatoknál a paraméterektől eltekintettünk, így a közölt adatok csak a származtatásnak tudhatók be.

A származtatott feladat megkonstruálásánál egy mohó algoritmust használtunk. Az algoritmus homogén szabás-vektorokat vizsgált csak, az i -edik idomnál a $C_1/L_1 \leq \dots \leq C_p/L_p$ sorrendben nézte végig az alapanyagokat. Az első olyat elfogadta, amelyikből ez az idom a kért darabszámban kiszabható (az aktuális raktárkészlet alapján). Ha nem volt ilyen alapanyag, akkor azt vette először, amelyikből a legnagyobb arányban kinyerhető, a fennmaradó darabszám alapján részrendelést hozott létre és újra kezdte a keresést stb. Általában 2–4 rendelést kellett felbontani 2–3 részrendelésre. A 4. táblázatból is látható, hogy az első fázis

4. táblázat. A származtatott /a/ és az eredeti /b/ feladatok futási jellemzői

	átlagos iterációs szám	átlagos futásidő /sec./	a futásidőből KP-kra /sec./
p=5 a.	44	198	148
m=15 b.	59	259	203
a/b %	75	76	73
p=10 a.	109	1.277	824
m=30 b.	127	2.095	1.668
a/b %	86	61	50

átugrása a megadott feltételek mellett jelentős futásidő csökkenést okoz. A kétfázisú módszernél minden esetben legalább m transzformáció kellett ahhoz, hogy eljussunk a megengedett bázismegoldáshoz. Az első csoportnál ($p = 5$ és $m = 15$) 2 feladatnál a második csoportnál ($p = 10$ és $m = 30$) 3 feladatnál a kétfázisú módszer igényelt kisebb futásidőt. A származtatott feladat memóriaigénye az eredetihez képest csak kevésbé nőtt (a feltételek száma 4–8-cal lett több), ezáltal a bázistranszformációk időigénye is alig növekedett. A hátizsákfeladatok megoldásának időszükségletét (amely az összes futásidő 65–80 %-át teszi ki) az idomok felbontása gyakorlatilag nem növeli, mivel a megegyező hosszúságú részrendelések közül csak egyet, a legkedvezőbb árnyékárút kell figyelembe venni.

5. A számítógépes alkalmazások néhány tapasztalata

A dolgozat lezárásakor a leírtakon alapuló programcsomagot 12 helyen alkalmazták, közülük 9 gépipari (fémrudakat és -lemezeket nyírnak), 2 bútorigari (bútorlapok szabása), 1 pedig építőipari vállalat. Az anyagleszabási problémák mindegyik üzemnél egyediek, a vezérlés hatékonysága nagyban függ a diszpécstől és a munkaszervezés színvonalától, ezért statisztikailag értékelhető információkról nem számolhatunk be. Tapasztalataink szerint az anyagleszabás tervezésének korábbi (főleg manuális) módszereihez képest kb. felére csökkent a hulladék. A vezérlések hulladékra gyakorolt hatása az előző fejezet mintapéldáihoz hasonló.

Az eltérő üzemi elvárások miatt minden egyes felhasználónál az adaptálás során ki kellett dolgozni egy alkalmazói stratégiát, amely általában egy paraméter (főleg a T_{uk}) következetes vezérlésen alapszik. Az alkalmazói stratégiát a felhasználó által megadott, néhány tipikusnak tekinthető feladat 3–4-szeri lefuttatásával ki lehet dolgozni.

A futásidőt jelentősen csökkenthetjük egy ε (> 0) küszöbszám megadásával. Ha a $3m$ -edik bázistranszformáció után valamely lépésnél a célfüggvény csökkenése ε -nál kisebb, akkor az addigi legjobb egészértékű megoldást optimálisnak fogadjuk el. Ezt a heurisztikus kiugrási feltételt alkalmazva a futásidő jelentősen csökkenthető, időnként közel felére, miközben a hulladék nagysága is elfogadható szinten van.

Megjegyezzük, hogy a modell áttekinthetőségét szoftver-eszközökkel is támogatjuk. A paraméterek egységesen megadhatók, illetve alapértelmezés is választható, az ismételt futtatáshoz megőrizzzük az adatokat. A kétdimenziós leszabások méretarányos rajza képernyőre kérhető, printeren és/vagy plotteren megjeleníthető stb. (bővebben [20]).

IRODALOM

- [1] ENG, G., DAELLENBACH, H.G., "Forest outturn optimization by Dantzig-Wolfe decomposition and dynamic programming column generation", *Operations Research* 33 (1985), 459–464.
- [2] DYCKHOFF, H., "A new linear programming approach to the cutting stock problem", *Operations Research* 29 (1981), 1092–1104.
- [3] DYCKHOFF, H. KRUSE, H.J., ABEL, D., GAL, T., "Trim loss and related problems", *Omega* 13 (1985), 59–72.

- [4] CHRISTOFIDES, N., WHITLOCK, C., "An algorithm for two-dimensional cutting problems", *Operations Research* 25 (1977), 30–44.
- [5] GALAMBOS, G., IMREH, B., "Egydimenziós szabási feladatok megoldása oszlopgenerálással", *Alk. Mat. Lapok* 25 (1984), 73–86.
- [6] GILMORE, P.C., "Cutting stock, linear programming, knapsacking, dynamic programming and integer programming, some interconnections", *Annals of Discrete Mathematics* 4 (1979), 217–235.
- [7] GILMORE, P.C., GOMORY, R.E., "A linear programming approach to the cutting stock problem", *Operations Research* 9 (1961), 849–859.
- [8] GILMORE, P.C., GOMORY, R.E., "A linear programming approach to the cutting stock problem—Part II", *Operations Research* 11 (1963), 863–888.
- [9] GILMORE, P.C., GOMORY, R.E., "Multistage cutting stock problems of two and more dimensions", *Operations Research* 13 (1965), 94–120.
- [10] GILMORE, P.C., GOMORY, R.E., "The theory and computation of knapsack functions", *Operations Research* 14 (1966), 1045–1074.
- [11] GOLDEN, B.L., "Approaches to the cutting stock problem", *AIIE Transactions* 8 (1976), 265–274.
- [12] HAESSLER, R.W., "Controlling cutting pattern changes in one-dimensional trim problems", *Operations Research* 23 (1975), 483–493.
- [13] HAESSLER, R.W., "A note on computational modifications to the Gilmore–Gomory cutting stock algorithm", *Operations Research* 28 (1980), 1001–1005.
- [14] HINXMAN, A.I., "The trim-loss and assortment problems: a survey", *EJOR* 5 (1980), 8–18.
- [15] KALMÁR, J., LENGYEL, I., "A raktárkészlet figyelembevétele a Gilmore–Gomory módszernél", *Sigma* 17 (1984), 167–183.
- [16] KALMÁR, J., LENGYEL, I., "A módosított szimplexmódszer első fázisának átugrása anyagleszabási feladatoknál", *MTA RKK* (1988), 14, (Kézirat).
- [17] LENGYEL, I. KUBA, A., "Síkűveg szabásának optimalizálása", *Sigma* 14 (1981), 169–190.
- [18] LENGYEL, I., "Darabolás — korszerű módszerekkel", *Ipargazdaság* 11 (1984), 19–31.
- [19] LENGYEL, I., "Anyagleszabási problémák hátizsákfeladatainak egy gyakorlati célú általánosítása", *Sigma* 19 (1986), 193–208.
- [20] LENGYEL, I., "Darabolási tervek készítése mikroszámítógépekkel", *Ipargazdaság* 6 (1988), 34–38.
- [21] MADSEN, OLI, B.G., "Glass cutting in a small firm", *Mathematical Programming* 17 (1979), 85–90.
- [22] TUCKER, A., "A structure theorem for the consecutive 1's property", *Journal of Combinatorial Theory* 12(B) (1972), 153–162.
- [23] VOUTILAINEN, R., "On the pattern sequencing problem — definitions, complexity and applications", *Paper presented at EURO IV., Cambridge* (1980), 28.
- [24] WANG, P.Y., "Two algorithms for constrained two-dimensional cutting stock problems", *Operations Research* 31 (1983), 573–586.

(Beérkezett: 1988. március 17.)

(Átdolgozva beérkezett: 1989. október 6.)

LENGYEL IMRE

MTA REGIONÁLIS KUTATÁSOK KÖZPONTJA ALFÖLDI KUTATÓCSOPORT

5600 BÉKÉSCSABA, SALLAI U. 40-42.

LAKÁS: 5600 BÉKÉSCSABA, WLASSICH SÉTÁNY 25. III. 17.

COMPUTATIONAL NOTES TO THE GILMORE--GOMORY
CUTTING STOCK ALGORITHM

I. LENGVEL

In this paper an extension and computational modifications of the Gilmore--Gomory cutting stock algorithm are presented so that constraints deriving from the practical circumstances of an industrial company can also be taken into account. A sample example of the one-dimensional problem of Gilmore and Gomory and that of the two-dimensional problem of Wang are solved, and their computational characteristics are being published. The computarization of the procedure has already been used by 12 industrial companies.

MEGJEGYZÉSEK A HEURISZTIKUS GRÁFKERESŐ ALGORITMUSOKHOZ

GREGORICS TIBOR

Budapest

A Mesterséges Intelligencia (MI) számos problémája megoldható egy olyan kereséssel, amikor a probléma lehetséges állapotai között a kezdőállapotból egy célállapotba vezető műveletsorozat megtalálása a feladat. A MI egyik jól ismert ilyen keresési technikája a *gráfkeresés*, amelynek talán a legismertebb osztálya az A^* algoritmus. Ebben a cikkben két olyan algoritmus osztályról és azok tulajdonságairól lesz szó, amelyek az A^* algoritmus általánosításai.

1. Bevezetés

A Mesterséges Intelligencia (MI) számos problémája megoldható egy olyan kereséssel, amikor a probléma lehetséges állapotai között a kezdőállapotból egy célállapotba vezető műveletsorozat megtalálása a feladat. Az állapotokkal és az azokon értelmezett véges számú műveletekkel reprezentált problémát megfeleltethetjük egy súlyozott irányított gráfnak, az úgynevezett *reprezentációs gráfnak*, ahol a csúcsok az állapotokat, az élek a (pozitív) költségeikkel együtt a műveletek által megvalósított átmeneteket szimbolizálják. Ebben a megoldásnak a kezdőcsúcsból egy célcsúcsba vezető irányított út felel meg, amelyet valamely gráfban történő kereső algoritmussal állítunk elő. A MI problémáira jellemző, hogy nagyon sok állapot között kell egy alkalmasat megkeresni, ezért a hagyományos, direkt algoritmusok használata gyakran kombinatorikus robbanáshoz vezet. A keresések akkor eredményesek, ha támaszkodhatnak a megoldandó feladat ismereteire. Ezért a sikeres és hatékony működés érdekében a MI algoritmusai különböző módon az adott feladattal kapcsolatos speciális tudást, heurisztikát építünk be.

A MI egyik jól ismert keresési technikája a *gráfkeresés*. Egy gráfkereső eljárás fokozatosan deríti fel az s startcsúcsból kivezető utakat, és ezeket a G keresőgráfban tartja nyilván. A felderített utak végén az úgynevezett *nyílt csúcsok* találhatók. (Ezek alkotják a *NYÍLT* halmazt.) Ezek közül az f *kiértékelő függvény* segítségével választjuk ki a legígéretesebbet. Az algoritmus mindig egy olyan csúcsot terjeszt ki, amelyekre ennek a függvénynek az értéke minimális. Ha több ilyen csúcs van, akkor a választás tetszőleges, illetve valamely más elv (ezt szokták *tie-breaking-rule*-nak nevezni) alapján dönt az eljárás. A kiválasztott nyílt csúcsot *kiterjesztjük*, azaz előállítjuk az utódait. A kiterjesztett csúcsokat *zárt csúcsoknak* nevezik. Az algoritmus akkor terminál, ha a kiválasztott csúcs egy célcsúcs, vagy már nincs folytatható út.

A keresések illusztrálására az egyik leggyakrabban alkalmazott probléma a 15-ös kirakó játék. Itt egy 4×4 -es keretben 15 darab lapocska és egy üres hely található. A feladat az, hogy egy megadott kiinduló konfigurációból elindulva kell elérni egy kívánt célkonfigurációt olyan mozgatásokkal, amelyek egy lapocskát a vele szomszédos üres helyre tolnak. A feladatot megoldhatjuk gráfkereséssel úgy, hogy nem használunk heurisztikát (szélességi keresés), vagy az úgynevezett rossz-helyek (W) illetve a távolság heurisztikát (P) használjuk. Az előbbi a nem célbeli helyükön álló lapocskák számával, az utóbbi a lapocskák célbeli helyétől vett *Hamilton-távolságok* összegével becsüli a hátralevő mozgatások számát. Az 1. ábrán egy konkrét kezdő és célkonfiguráció mellett a különböző heurisztikákat használó gráfkeresések futási eredményét láthatjuk. A heurisztikus keresések megoldják a problémát, a heurisztikát nem használó keresés viszont nem. (1. ábra)

Alkalmazott heurisztika	\emptyset	W	P
A talált megoldás hossza	Out of memory	21	21
Futási idő	Out of memory	4' 1.84''	0' 1.86''

1. ábra

Az A^* algoritmussal talán a legismertebb heurisztikus gráfkeresés, amelynek kiértékelő függvénye minden egyes nyílt csúshoz hozzárendeli a startcsúsból odavezető nyilvántartott út költségének és az úgynevezett heurisztikus függvényértéknek összegét. Ez utóbbi a reprezentációs gráf csúcsainál a legolcsóbb hátralevő út költségére ad alsó becslést. Ilyenkor szokták ezt a függvényt megengedhetőnek (*admissible*) nevezni. Az A^* algoritmus a megoldás létezése esetén megtalálja az optimális megoldást. Az előző példában a W és a P heurisztika alsóbecslést adott, így a talált megoldás optimális.

Ebben a cikkben két olyan algoritmus osztályról és azok tulajdonságairól lesz szó, amelyek az A^* algoritmus általánosításai.

Először azokat az általános gráfkereső algoritmusra vonatkozó fontosabb jelöléseket és eredményeket ismertetjük, amelyekre a későbbiek során szükségünk lesz.

Ezután az úgynevezett *A algoritmussal* foglalkozunk. Ez abban az értelemben általánosítása az A^* algoritmusnak, hogy nála nem követeljük meg a heurisztika megengedhetőségét. A szakirodalomban gyakran összemosiszódik az A és az A^* algoritmus fogalma, hiszen csak a felhasznált heurisztikus függvénytől függ az, hogy melyik algoritmusról beszélünk. A két algoritmus szétválasztásával csak néhány esetben találkozhatunk [9], [8]. Most be fogjuk bizonyítani, hogy ez az algoritmus a megoldás létezése esetén mindig talál egy megoldást, még végtelen nagy gráfok esetén is. Ha figyelembe vesszük, hogy sok esetben egy nem megengedhető heurisztikát használó A algoritmus hatékonyabb, mint egy A^* algoritmus, akkor ennek

a tulajdonságnak a jelentősége nem is kicsi. Sokszor ugyanis a megoldó algoritmus hatékonysága (elsősorban a gyorsasága) előbbre való szempont, mint az, hogy optimális megoldást találjunk. Egy *Rubik-kocka* kirakó versenyen sem az győz, aki hosszadalmas gondolkodás árán ugyan, de a legkevesebb forgatással éri el a kitűzött alakzatot (bár teljesítménye csodálatra méltó), hanem az, akinek az összideje (gondolkodás és forgatások) a legrövidebb. Az optimális megoldás megtalálása ugyanis általában hosszabb futási időt igényel.

Ezt szemlélteti a 15-ös kirakó játék esetén a rossz-helyek (W), a távolság (P) és a keret ($4S + W$) heurisztika összehasonlítása is. A keret heurisztika első komponense (S) egy adott konfigurációban egyrészt azt számolja, hogy a széleken az órajárásával megegyező irányban egy lapocskát hányszor nem a célbeli szomszédja követi, valamint azt, hogy hány sarokban nem a célbeli lapocskára található. Ez a heurisztika nem megengedhető ugyan, és így nem is garantálja az optimális megoldás megtalálását, de az esetek többségében lényegesen hatékonyabb algoritmushoz jutunk általa. A gráfkereső algoritmus egy olyan kezdő- illetve célállapotra, ahol az optimális távolság 30 lépés volt, a rossz-helyek heurisztika mellett „out of memory” üzenettel állt le, a távolság heuristikával 31.86 másodperc múlva, míg a keret heurisztika 4.23 másodperc után egy 58 lépés hosszú megoldással (2. ábra).

Alkalmazott heurisztika	W	P	$4S + W$
A talált megoldás hossza	Out of memory	30	58
Futási idő	Out of memory	31.86"	4.23"

2. ábra

Másodjára egy olyan algoritmus osztályt (AB) vezetünk be, amely egyfelől tartalmazza az A^* algoritmust, másfelől az úgynevezett B algoritmust is, amely köztudottan egy hatékonyabb változata az A^* algoritmusnak. Ezen algoritmus osztály lehetővé teszi a B algoritmus fenti tulajdonságának egyszerű bizonyítását.

2. Az általános gráfkereső algoritmus és tulajdonságai

Először megadjuk a gráfkeresés általános algoritmusát, amely alkalmas arra, hogy belőle származtathatjuk a szakirodalomban előforduló fontosabb gráfkereső algoritmusokat (mélységi, szélességi, egyenletes, A^* stb.). Ennek módja az, hogy az algoritmus által használt kiértékelő függvényt kell alkalmasan megválasztanunk.

Az algoritmusban g -vel jelöltük azt a költségfüggvényt, amely az egyes csúcsoknál nyilvántartott startcsúcsból odavezető út költségét adja meg, p -vel azt

a visszafelé mutató pointert, amelyik minden csúcshoz annak pontosan egy szülőjét megadva egyértelműen kijelöl egy feszítőfát a mindenkor keresőgráfon. A $G := G \cup \Gamma(n)$ jelöli azt, amikor egy n csúcs kiterjesztése árán új csúcsokkal és élekkel bővül a keresőgráf.

Procedure GRÁFKERESÉS

```

1.  $G \leftarrow \{s\}$ ; NYÍLT  $\leftarrow \{s\}$ ;  $g(s) \leftarrow 0$ ;  $p(s) \leftarrow \text{nil}$ ;
2.  $n \leftarrow s$ ;
3. while not(üres(NYÍLT) or cél( $n$ )) loop
4.    $M \leftarrow \Gamma(n)$ ;
5.   while not üres( $M$ ) loop
6.      $m \leftarrow \text{elem}(M)$ ;  $M \leftarrow M \setminus \{m\}$ ;
7.     if  $m \notin G$  or  $g(n) + c(n, m) < g(m)$  then
8.        $g(m) \leftarrow g(n) + c(n, m)$ ;  $p(m) \leftarrow n$ ;
9.        $\text{NYÍLT} \leftarrow \text{NYÍLT} \cup \{m\}$ ;
10.    endif
11.  endloop
12.   $G \leftarrow G \cup \Gamma(n)$ ;  $\text{NYÍLT} \leftarrow \text{NYÍLT} \setminus \{n\}$ ;
13.   $n \leftarrow \min_f(\text{NYÍLT})$ ;
14. endloop
end

```

Az általános gráfkereső algoritmus véges reprezentációs gráfokon mindig terminál [8]. A következő lemma az algoritmus invariáns állítását fogalmazza meg. Ez a lemma a kulcsa a gráfkereső algoritmusokkal kapcsolatos állítások bizonyításainak.

2.1. LEMMA. Legyen n a reprezentációs gráf egy olyan tetszőleges csúcsa, amelyhez vezet út a startcsúcsból. Az általános gráfkereső algoritmus az n csúcs kiterjesztése előtt bármelyik $s \xrightarrow{*} n$ optimális úton nyilvántart egy olyan m csúcsot, amelyikre teljesül, hogy

- $$\begin{aligned}
 (1) \quad & m \in \text{NYÍLT} \\
 (2) \quad & g(m) = g^*(m)
 \end{aligned}$$

Bizonyítás. Rögzítsünk egy tetszőleges n csúcsot, amelyhez vezet út a startcsúcsból. Megmutatjuk, hogy bármely $s \xrightarrow{*} n$ optimális úton van olyan csúcs, amelyre teljesülnek a tétel (1) (2) állításai.

Bizonyítsunk az algoritmus által végzett kiterjesztések száma szerinti teljes indukcióval.

A legelső kiterjesztés előtt egyetlen eleme van a NYÍLT halmaznak, ez a startcsúcs. A startcsúcs minden s -ből kiinduló útnak az első eleme, és így rajta van minden $s \xrightarrow{*} n$ optimális úton is. A $g(s) = 0$ miatt nyilván teljesül, hogy $g(s) = g^*(s)$.

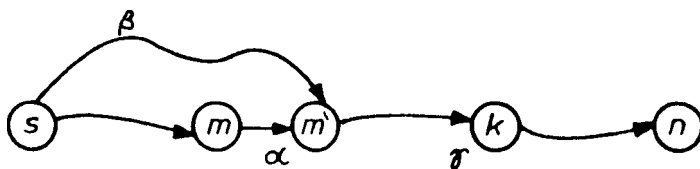
Tegyük fel, hogy az i -edik kiterjesztéskor teljesül az állítás. Jelölje m csúcs a tetszőlegesen rögzített $\alpha \in \{s \xrightarrow{*} n\}$ optimális úton elhelyezkedő (1) (2) tulajdonságú csúcspot. Lássuk be, hogy az $i + 1$ -edik kiterjesztés után is lesz egy ilyen tulajdonságú csúcs az α úton.

Ha a soron következő kiterjesztés nem az m csúcspot veszi ki a *NYÍLT* halmazból, akkor az m csúcs továbbra is az állításnak megfelelő csúcs. Ha viszont az m csúcspot terjeszti ki az algoritmus, akkor annak utódai között szerepel az α úton elhelyezkedő m' csúcs, és mivel α optimális,

$$g(m') = g^*(m) + c(m, m') = g^*(m').$$

Az m' csúcs csak abban az esetben nem lesz az $i + 1$ -edik kiterjesztés után *NYÍLT* halmazban, ha már korábban — az m csúcs kiterjesztése előtt — talált az algoritmus az m' -höz egy az m -et elkerülő ($g^*(m')$ költségű) optimális utat, amelynek minden ℓ csúcsa végre lezárult $g(\ell) = g^*(\ell)$ költség mellett.

Jelöljük ezt az utat β -val, az α út $m' \rightarrow n$ szakaszát pedig γ -val (3. ábra). Tekintsük a β és a γ konkatenáltját. Ez a $\beta\gamma$ út nyilván optimális út a startcsúcs és az n csúcs között, így az indukciós feltevés miatt tartalmaz egy (1) (2) tulajdonságú k csúcspot az i -edik kiterjesztéskor.



3. ábra

Másrészt ez a k csúcs biztosan a γ szakaszon helyezkedik el, hiszen β -nak nincs nyílt csúcsa, így a k csúcs az α útnak is eleme. A k csúcs az m csúcs kiterjesztése után is ott marad a *NYÍLT* halmazban, azaz van az $i + 1$ -edik lépésben is megfelelő tulajdonságú csúcs az α úton. \square

2.1. TÉTEL. Ha egy véges reprezentációs gráfban létezik megoldás, akkor az általános gráfkereső algoritmus egy célcsúcs megtalálásával terminál.

Bizonyítás. Mivel a gráfkereső algoritmus véges gráfokon mindig terminál, így elegendő azt megmutatnunk, hogy megoldás létezése esetén a *NYÍLT* halmaz nem ürülhet ki. Vagyis ha az algoritmus terminál, akkor azt csak egy célcsúcs megtalálásával teheti.

Ha a 2.1. Lemma állításában szereplő még ki nem terjesztett csúcsnak egy olyan célcsúcsot választunk, amelyhez vezet út a startcsúcsból, akkor a célcsúcs kiválasztása (megtalálása) előtt a hozzávezető úton mindig van legalább egy olyan csúcs, amely eleme a *NYÍLT* halmaznak, amely így nem válhat üressé. \square

3. Az *A* algoritmus

3.1. Definíció. *A* algoritmusnak nevezzük azt az általános gráfkereső algoritmust, amelynek kiértékelő függvénye

$$f(n) = g(n) + h(n) \quad \forall n \in \text{NYÍLT}$$

és ahol

$$h : N \rightarrow \mathbb{R}$$

$$h(n) \geq 0 \quad \forall n \in N.$$

3.1. LEMMA. Ha az *A* algoritmus nem terminál, akkor minden *NYÍLT* halmazba került csúcs véges sok lépés után kiterjesztésre kerül.

Bizonyítás. Legyen *m* egy nyílt csúcs. Elég megmutatni, hogy azon csúcsok száma véges, amelyeknek kiértékelő függvényértéke kisebb vagy egyenlő az *f(m)*-nél. Mivel egy csúcsot legfeljebb véges sokszor terjeszt ki az algoritmus, így véges lépésen belül bekövetkezik, hogy a *NYÍLT* halmazban az *f(m)* értékénél csak nagyobb függvényértékű csúcsok fognak szerepelni, azaz az *m* csúcs kiterjesztésére sor kerül.

Legyen *n* egy tetszőleges csúcsa a reprezentációs gráfnak. Amikor ez a csúcs bekerül a *NYÍLT* halmazba, akkor fennáll, hogy

$$f(n) = g(n) + h(n) \geq g^*(n) \geq d^*(n) * \delta,$$

ahol $d^*(n)$ a legrövidebb $s \rightarrow n$ út hosszát, δ pedig az élek költségének alsó korlátját jelöli. A fenti becslés azt mutatja, hogy egy csúcs kiértékelő függvényértéke arányos a csúcs mélységével.

Ha az *m* egy adott nyílt csúcs *f(m)* rögzített függvényérték mellett, akkor legyen a továbbiakban *d*-vel jelölt mélység az $f(m)/\delta$ felső egészértéke. A *d*-nél mélyebben elhelyezkedő *k* csúcsokra

$$f(k) \geq d^*(k) * \delta > d * \delta \geq f(m)$$

teljesül. Ebből az következik, hogy minden olyan *l* csúcs, amelyre

$$f(l) \leq f(m)$$

fennáll, a *d* mélység felett helyezkedik el. A reprezentációs gráf tulajdonságai (minden csúcsból véges sok él vezet ki) miatt ilyen csúcsból csak véges sok van. \square

3.2. TÉTEL. Az A algoritmus mindig talál megoldást feltéve, hogy létezik.

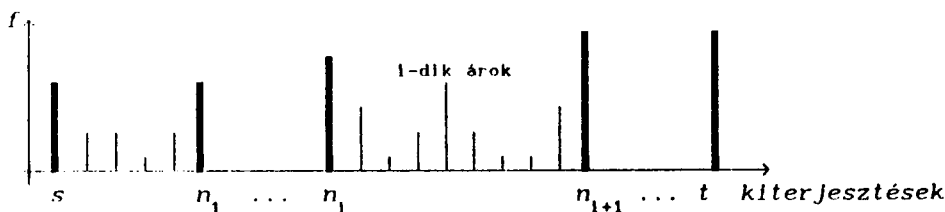
Bizonyítás. Véges reprezentációs gráfokra az állítás a 2.1. tételből következik. Elegendő tehát olyan végtelen reprezentációs gráfokon meggondolni az állítást, amelyben a feltétel szerint vezet út az s -ből a T -be.

Legyen $\alpha \in \{s \xrightarrow{*} t\}$ egy optimális út, ahol $t \in T$ egy célcsőcs. A 2.1. lemma szerint mindig van ezen az úton egy nyílt csűcs. Vegyük a legelső ilyet. Ha az eljárás nem terminálna, akkor ez a nyílt csűcs a 2.1. lemma következtében véges lépésen belül végleg kiterjesztésre kerül, és ekkor az α úton fekvő valamelyik leszármazottja lesz a 2.1. lemmának megfelelő nyílt csűcs. E gondolatmenetet tovább vezetve véges lépés után a t célcsőcs kiválasztására kerül sor, azaz az eljárás terminál. Erre csak abban az esetben nem kerül sor, ha a keresés már korábban, egy másik célcsőcs elérésével terminál. \square

E fenti eredmény jól használható olyan feladatok megoldásánál, amikor nem fontos optimális megoldást adnunk, de egyfelől tudjuk, hogy van megoldás, másfelől az olyan keresések, mint a mélységi vagy szélességi, nem működnek hatékonyan. Konkrét futási eredmények igazolják azt, hogy egy nem megengedhető, de ötletes heurisztikával jóval hatékonyabb algoritmushoz jutunk, mintha optimális megoldást garantáló A^* algoritmust használnánk.

4. Az AB algoritmus osztály

Vegyük szemügyre az A^* algoritmus működését. Ábrázoljuk egy grafikonon a kiterjesztésre kiválasztott csűcsokhoz tartozó kiértékelő függvényértékeket (4. ábra).



4. ábra

Először az s startcsűcs kerül kiterjesztésre $h(s)$ nagyságú függvényértékkel. Ezután egy csűcs többször is kiterjesztésre kerülhet, de annak újabb kiterjesztéséhez egyre kisebb kiértékelő függvényérték tartozik. Kiválasztható a grafikonon látható értékek közül egy monoton növekvő részsorozat a következőképpen. Vegyük

bele a legelső értéket, majd mindig a soronkövetkező legközelebbi olyat, amelyik nagyobb vagy egyenlő minden korábbi értéknél. A kiválasztott értékeket a továbbiakban küszöbértékeknek, a hozzájuk tartozó csúcsokat küszöbcsúcsoknak fogjuk nevezni. Egy csúcs csak az első kiterjesztésekor lehet küszöbcsúcs, a további kiterjesztéseire két küszöbcsúcs közötti „árokban” kerülhet sor, azaz minden küszöbcsúcs különböző.

Ha feltesszük, hogy létezik megoldás, akkor az utolsó küszöbcsúcs nyilván a célcsúcs lesz, ugyanis ennek értékénél minden korábbi kiterjesztésnél szereplő kiértékelő függvényérték kisebb vagy egyenlő. Ez egy fontos tulajdonsága az A^* algoritmusoknak [8].

Az algoritmus a küszöbcsúcsok közötti árokban is kiterjeszthet új csúcsot, de gyakran kerül itt sor egy már kiterjesztett, azaz zárt csúcs újbóli kiterjesztésére. Ha megvizsgáljuk, hogyan működik egy kevésbé hatékony, azaz túl sok kiterjesztést végző A^* algoritmus, azt tapasztaljuk, hogy az egyes küszöbcsúcsok közötti árok túl hosszúak, ugyanabban az árokban egy csúcsot többször is kiterjeszt az algoritmus. Jelentősen javíthatnánk az algoritmus hatékonyságát, ha sikerülne úgy módosítani azt, hogy egy csúcs egyazon árokban legfeljebb egyszer kerüljön kiterjesztésre. Ezt úgy érhetjük el, hogy nem az f kiértékelő függvény alapján, hanem más, megfelelően megválasztott sorrend szerint terjesztjük ki a csúcsokat két szomszédos küszöbcsúcs között.

Definiáljunk most egy új algoritmus osztályt.

4.1. Definíció. Az AB algoritmusosztályt az általános gráfkereső algoritmusból származtatjuk úgy, hogy a kiterjesztésre történő kiválasztás stratégiáját megváltoztatjuk. Ez az algoritmus 13. lépésének módosításával jár.

```

if  $\min_f (NY\acute{I}LT) < F$  then
     $n \leftarrow \min_q (m \in NY\acute{I}LT \mid f(m) < F)$ 
else
     $n \leftarrow \min_f (NY\acute{I}LT);$ 
     $F \leftarrow f(n)$ 
endif

```

ahol $q : NY\acute{I}LT \rightarrow \mathbb{R}$ egy tetszőleges függvény, az f továbbra is $g + h$ alakú függvény, F pedig a legutolsó küszöbérték. Ez utóbbinak kezdőértékét az algoritmus 2. lépésében az

$$F \leftarrow f(s)$$

értékadással kell adni.

Az AB algoritmusban alkalmazott q függvény biztosítja azt az új stratégiát, amelyre két küszöbcsúcs között váltunk át. Nyilvánvaló, hogy $q = f = g + h$ választás mellett az AB algoritmus éppen az A algoritmus lesz, megengedhető heurisztika mellett pedig az A^* algoritmus.

4.1. TÉTEL. Az AB algoritmus osztály minden azonos módon működő (ugyanazt a tie-breaking-rule-t és heurisztikát használó) tagja a q függvénytől függetlenül ugyanazokat a csúcsokat terjesztik ki; a küszöbcsúcsok és küszöbértékeik megegyeznek, a megfelelő küszöbcsúcsok kiterjesztésekor a nyilvántartott keresőgráfjaik, feszítőfájuk, nyílt halmazaik azonosak.

Bizonyítás. Jelöljük G_i -vel az i -edik küszöbcsúcs kiterjesztésekor meglévő keresőgráfot, és $NYÍLT_i$ -vel az abban levő nyílt csúcsok halmazát. Belátjuk, hogy a G_i , és a $NYÍLT_i$ a q függvénytől függetlenül minden i -re megegyezik az azonosan működő AB algoritmusoknál.

Az állítást a küszöbcsúcsok száma szerinti teljes indukcióval bizonyítjuk. $i = 0$ -ra nyilvánvalóan igaz az állítás, hiszen ekkor minden AB algoritmus csak a startcsúcsot ismeri. Tegyük fel, hogy az i -edik küszöbcsúcs kiterjesztésekor igaz az állítás. Be kell látni, hogy az $i + 1$ -edik küszöbcsúcs kiterjesztésekor is teljesül.

Ehhez mindössze azt kell megmutatni, hogy q -tól függetlenül ugyanazon csúcsokat terjeszti ki az AB algoritmus az i -edik küszöbcsúcs utáni árokban.

Jelölje F_i az i -edik küszöbértéket, és $g_i(n)$ az $n \in G_i$ csúcsoknál az i -edik küszöbcsúcs kiterjesztésekor nyilvántartott költségértéket. Tekintsük a következő halmazt:

$$\begin{aligned} D_i = \{n_i\} \cup \{m \in N \mid & (1) \exists \alpha \in \{n_i \rightarrow m\} \text{ úgy, hogy } \forall n \in \alpha \setminus \{m\} : n \in D_i, \\ & (2) g_i(n_i) + k^\alpha(n_i, m) < g_i(m), \text{ ha } m \in G_i, \\ & (3) g_i(n_i) + k^\alpha(n_i, m) + h(m) < F_i\}. \end{aligned}$$

A D_i egyrészt nyilvánvalóan független a q függvénytől, másrészt az AB algoritmusok pont a D_i -hez tartozó csúcsokat terjesztik ki az i -edik árokban. Ugyanis ezek a csúcsok az n_i kiterjesztése után biztos felkerülnek az AB algoritmusok $NYÍLT$ halmazába az (1) és (2) tulajdonságok miatt, a (3) tulajdonság szerint pedig még a következő küszöbcsúcs előtt kiterjesztésre kerülnek. Egyéb csúcsot pedig nem terjeszthet ki itt az algoritmus. \square

A fenti tételnek következménye, hogy az AB algoritmusok megengedhető heurisztika mellett ugyanazzal az optimális megoldással terminálnak, mint az A^* algoritmus feltéve, hogy létezik a megoldás.

A 4.1. tétel nem megengedhető heurisztika esetén is igaz. Az AB algoritmusok ilyenkor is mindig terminálnak, akárcsak az A algoritmus, bár nem feltétlenül ugyanazzal a megoldással, hiszen ilyenkor az utolsó küszöbcsúcs után van egy befejezetlen, nem küszöbcsúccsal végződő árok, amelyen belül a csúcsok kiterjesztésének sorrendje változó az egyes AB algoritmusoknál.

Az egyes AB algoritmusok közti egyetlen különbség abban áll, hogy az egyes árkokban más sorrendben terjesztik ki a csúcsokat, és ebből adódóan előfordulhat, hogy valamely csúcsot ugyanazon árkon belül az egyik algoritmus többször, a másik kevesebbszer terjeszt ki. A következő definíció az AB algoritmusok azon tagját vezeti be, amely legfeljebb egyszer terjeszt ki egyazon árkon belül egy csúcsot.

B algoritmusnak nevezzük azt az AB algoritmust, amely q függvénye az algoritmus által számolt g költségfüggvénnyel egyezik meg [6].

4.2. TÉTEL. A B algoritmus egy árkon belül egy csúcsot legfeljebb egyszer terjeszt ki.

Bizonyítás. Tegyük fel indirekt, hogy van egy olyan $i \in \mathbb{N}$, hogy az i -edik árokban egy $m \in D_i$ csúcsot a B algoritmus kétszer terjeszt ki. Ez azt jelenti, hogy az algoritmus először egy $\alpha \in \{n_i \rightarrow m\}$ D_i -beli út mentén terjeszti ki az m csúcsot ($g(m) = g_i(n_i) + k^\alpha(n, m)$), ahol $g_i(n_i)$ az n_i kiterjesztésekor az ott nyilvántartott költség, majd talál egy olcsóbb $\beta \in \{n_i \rightarrow m\}$ D_i -beli utat. Nyilván

$$k^\beta(n_i, m) < k^\alpha(n_i, m)$$

Jelölje m' azt az elemét a β útnak, amely a $NYÍLT$ halmazban van az m csúcs α út mentén történő kiterjesztésekor. Mivel ekkor az algoritmus g függvény szerint választotta az m csúcsot kiterjesztésre, így

$$g(m) \leq g(m')$$

Ez azonban ellentmond annak, hogy

$$g(m') \leq g_i(n_i) + k^\beta(n_i, m') \leq g_i(n_i) + k^\beta(n_i, m) < < g_i(n_i) + k^\alpha(n_i, m) = g(m). \quad \square$$

Természetesen a tétel állítása nem zárja ki azt, hogy egy csúcs nem kerülhet újból kiterjesztésre egy későbbi árokban.

A B algoritmus tehát az AB algoritmosztálynak az a tagja, amelyik a lehető legkevesebb kiterjesztést végzi, miközben a többi tulajdonsága megengedhető heurisztika mellett az A^* algoritmusával azonos.

IRODALOM

- [1] BARR, A., FEIGENBAUM, E. A., *The Handbook of Artificial Intelligence I* (HeurisTECH Press, Stanford, 1982).
- [2] FEKETE, I., GREGORICS, T., VARGA, L. ZS., „Corrections to Graph-searching Algorithms”, *Procds: „Fourth Conference of Program Designers”, Budapest, June 1-3, 1988, (Ed. by A. Iványi) (ELTE, Budapest, 1988), 25-30.*
- [3] GREGORICS, T., „Another Introduce to Consistent Algorithms”, *Procds: „First Seminar on Artificial Intelligence”, Visegrád, January 23-24, 1989, (Ed. by I. Fekete, S. Nagy) (ELTE, Budapest, 1989), 137-144.*
- [4] GREGORICS T., „Which Graphsearch Algorithm is Better?”, *Procds: „Second Conference on Artificial Intelligence”, Budapest, January 23-25, 1991, (Ed. by I. Fekete, P. Koch), vol. 2 (ELTE, Budapest, 1991), 275-284.*
- [5] HART, P., NILSSON, N. J., RAPHAEL, B., „A Formal Basic for the Heuristic Determination of Minimum Cost Paths”, *IEEE Trans. System, Man and Cybernet* 4 (1968).
- [6] MARTINELLI, A., „On the Complexity of Admissible Search Algorithms”, *Artificial Intelligence* 8, No. 1 (1977).

- [7] MÉRŐ, L., „A Heuristic Search Algorithm with Modifiable Estimate”, *Artificial Intelligence* 23, No. 1 (1984).
- [8] NILSSON, N. J., *Principles of Artificial Intelligence* (Springer-Verlag, 1982).
- [9] PEARL, J., *Heuristics* (Reading, Mass, Addison-Wesley, 1984).
- [10] RICH, E., *Artificial Intelligence* (MacGraw-Hill Book Company, 1983).

(Beérkezett: 1992. augusztus 28.)

GREGORICS TIBOR
ELTE TTK ÁLTALÁNOS SZÁMÍTÁSTUDOMÁNYI TANSZÉK
1117 BUDAPEST, BOGDÁNFY ÚT 10/B

REMARKS TO THE HEURISTIC GRAPH-SEARCH ALGORITHMS

T. GREGORICS

A number of problems in the Artificial Intelligence (AI) area can be related to the general problem of finding a path through a space of problem states from the initial state to any goal state. One of the search techniques solving AI problems is the graphsearch. The most famous algorithm of these is the algorithm A^* . In this paper two algorithms and their properties are going to be discussed which are the generalizations of the algorithm A^* .

ANYAGMEGMARADÁS ÖSSZETETT KÉMIAI MECHANIZMUSOKBAN

DEÁK JENŐ, TÓTH JÁNOS ÉS VIZVÁRI BÉLA

Budapest, Gödöllő, Ankara

A jelen dolgozat célja, hogy az anyagmegmaradásra vonatkozó vizsgálatokat egységes elméletként mutassa be az Olvasónak. E célból bizonyításokkal és példákkal együtt összefoglalja az összetett kémiai mechanizmusok konzervativitására vonatkozó irodalmi eredményeket, másrészt az ismerteknél jobban kezelhető szükséges, elégséges és ekvivalens feltételeket, illetve algoritmusokat ad meg a konzervativitásra. Az eredmények nemcsak a konzervativitásra (és a vele rokon szub- és szuperkonzervativitásra és hasonló fogalmakra) vonatkoznak, hanem egy új, az eddigiekhez képest duális fogalomrendszerre is, amely az anyagmegmaradást (és a vele rokon fogalmakat) nem a kémiai komponensek anyagmennyiségének követésével, hanem az elemi reakciók hatásának figyelésével írja le. A két fogalomrendszer összehasonlításának legfontosabb eszköze egy általános, lineáris egyenlőtlenségekre vonatkozó alternatíva-tétel.

0. Bevezetés	74
1. Előkészületek	75
1.1. Elnevezések	75
1.2. Jelölések	76
a) Általános jelölések	76
b) A szövegben bevezetett speciális jelölések	76
2. Alapfogalmak	76
2.1. Mechanizmus, tömegmegőrzés, konzervativitás	77
2.2. Atomos szerkezetű specieszeket tartalmazó mechanizmusok	82
2.3. Rekeszrendszerek és elsőrendű mechanizmusok	83
2.4. A Volpert-gráf	84
3. A definiált tulajdonságok fennállására vonatkozó szükséges, elégséges, valamint szükséges és elégséges feltételek	85
3.1. A 2.2. definícióban szereplő fogalmakra vonatkozó feltételek	85
3.2. Konzervativitásra vonatkozó szükséges és elégséges feltételek	85
3.3. Az aciklikusság következményei	86
3.4. A rekeszrendszerek különféle általánosításai	88
3.5. Az atomszám-megmaradás törvénye	91
3.6. Kiegészítések	91
4. Kapcsolatok a definiált tulajdonságok között	92
5. Kitekintés	95
Függelék	95
Irodalom	96

0. Bevezetés

Dolgozatunk kettős céllal íródott. Egyrészt egységes keretben, bizonyításokkal és példákkal együtt összefoglaljuk az irodalomban fellelhető olyan eredményeket, amelyek az anyagmegmaradásra vonatkoznak, és kémiai reakciókinetikai indíttatásúak. Másrészt megadunk néhány olyan feltételt, illetve algoritmust, amelyek az ismertekkel szemben bizonyos előnyökkel rendelkeznek.

Hogyan alkalmazhatók az itt előadott állítások? Ha a vizsgált mechanizmus közvetlenül egy jól kézbentartott kísérlet leírására szolgál, akkor valószínűleg atomos szerkezetű kémiai komponenseket tartalmaz, és teljesíti az atomszámmegmaradás törvényét, tehát konzervatív. Ha viszont (és elsősorban erre a helyzetre gondoltunk) valaki konstruál egy mechanizmust, mert úgy gondolja, hogy az érdekes tulajdonságokkal rendelkezik (így született többek között a Brüsszelátor, a Schlögl-modell vagy az Ivanova-modell), akkor felvetődik az a kérdés, hogy a modell mennyire reális abban az értelemben, hogy a tömegmegmaradás alapvető követelményének eleget tesz-e. Erre a kérdésre, s az általa indukált rokon problémákra igyekeztünk ebben a dolgozatban válaszolni. Kiemeljük, hogy a szubkonzervativitásnak messzemenő *dinamikai* következményei vannak [10] a reakciók *determinisztikus* modelljére nézve, tehát igen hasznos, ha rendelkezünk erre vonatkozó feltételekkel. A kémiai reakciók szokásos *sztochasztikus* modelljében pedig a *stacionárius* eloszlásra vonatkozó állításokat lehet megfogalmazni az itt kifejtett állításokból kiindulva.

A cikk felépítése a következő: először ismertetjük a formális reakciókinetikából felhasználandó *alapfogalmakat*. Kiemeljük, hogy ezen fogalmak közül újak azok, amelyek a 2.2. definícióban jelennek meg. Ezután áttérünk az *eredmények* ismertetésére, ahol a tételek inkább logikai, mint történeti sorrendben követik egymást. Ismert tételek esetén mindenütt megadjuk az eredeti lelőhely pontos adatait. A Kitekintésben rámutatunk arra, hogy az itteni eredmények hogyan kapcsolódnak konkrét modellekre (elsősorban a szokásos sztochasztikus modellre) vonatkozó vizsgálatokhoz. Megemlítünk néhány nyílt problémát is. A Függelékben segédeszközként használt tételeket ismertetünk.

A további kapcsolatok iránt érdeklődő Olvasónak az [1,7,8] írásokat ajánljuk szíves figyelmébe.

A félreértések elkerülése végett megadjuk azon jelölések jegyzékét, amelyek talán nem szerepelnek mindenütt teljesen egységesen.

1. Előkészületek

1.1. Elnevezések

(Igen sok elnevezést vezettünk be közvetlenül az alapvető 2.1. definíció után; ezek mellett is a 2.1. szám szerepel.)

Név	A nevet tartalmazó definíció száma
atom	2.4
atommátrix	2.4
atomos szerkezetű komponenseket tartalmazó mechanizmus	2.4
atomszám-megmaradás törvénye	2.5
általános rekeszrendszer	2.7
elemi reakció	2.1
elemi reakció rendje	2.8
előrendű mechanizmus	2.8
kémiai komponens	2.1
kompartment-rendszer	2.6
komplex	2.1
komplex vektor	2.1
konzervatív	2.3
kúp	2.1
mechanizmus	2.1
mechanizmus rendje	2.8
reakcószimplex	2.1
reaktáns komplex	2.1
reaktáns komplex vektor	2.1
rekeszrendszer	2.6
reverzibilis elemi reakció	2.1
reverzibilis mechanizmus	2.1
speciesz	2.1
szigorúan szubkonzervatív	2.3
szigorúan szuperkonzervatív	2.3
szigorúan tömegfogyasztó	2.2
szigorúan tömegtermelő	2.2
stöchiometriai együttható	2.1
stöchiometriai tér	2.1
szubkonzervatív	2.3
szuperkonzervatív	2.3
tartó	2.1
termék komplex	2.1
termék komplex vektor	2.1
tömegfogyasztó	2.2
tömegmegőrző	2.2
tömegtermelő	2.2
üres komplex	2.1
Volpert-gráf	2.9
zárt általánosított rekeszrendszer	2.7
zárt rekeszrendszer	2.6

1.2. Jelölések

a) Általános jelölések

e_r	a standard bázis r -edik egységvektora
I_M	$M \times M$ -es egységmátrix
\mathcal{H}	a \mathcal{H} halmaz számossága
\mathbb{N}	a pozitív egész számok halmaza
\mathbb{N}_0	a nemnegatív egész számok halmaza
\mathbb{R}^+	a pozitív valós számok halmaza
\mathbb{R}^-	a negatív valós számok halmaza
\mathbb{R}_0^+	a nemnegatív valós számok halmaza
\mathbb{R}_0^-	a nempozitív valós számok halmaza
$\text{rank}(\gamma)$	a γ mátrix rangja
0	a nulla mátrixot jelöli (ami a szövegtől függően vektorrá specializálódhat)
1	összegező vektor; minden koordinátája 1
\top	a transzponált jele
\perp	az ortogonális komplementer jele
$<, \leq, >, \geq$	vektorokra és mátrixokra koordinátánként értendő

b) A szövegben bevezetett speciális jelölések

$\mathcal{A}(d)$	a d -edik atom
\mathcal{D}	az atomok halmaza
$\mathcal{K}(\gamma)$	a γ mátrix oszlopai által generált kúp
$\ell(\eta)$	az η komplex hossza
\mathcal{M}	a kémiai komponensek (speciesszek) halmaza
\mathcal{M}_0	a Volpert-gráf kezdőpontjainak halmaza
\mathcal{R}	az elemi reakciók halmaza
$\mathcal{S}(\gamma)$	a γ mátrix oszlopai által kifesztett lineáris altér
$\text{supp } \eta$	az η komplex tartója
$\mathcal{X}(m)$	az m -edik kémiai komponens vagy speciesz
α, β	a sztöchiometriai együtthatók mátrixai
δ	az atommátrix
\vdash	két vektor vagy mátrix között fönnáll, ha \leq és \neq egyaránt teljesül rájuk
\vdash	két vektor vagy mátrix között fönnáll, ha \geq és \neq egyaránt teljesül rájuk
\emptyset	üres komplex

A vektorok jelölésére félkövér kisbetűket, mátrixokéra félkövér latin nagybetűket (esetenként félkövér görög kisbetűket) használunk; a vektorok, illetve a mátrixok komponenseit pedig a megfelelő vékony, illetve dőlt betűvel jelöljük.

2. Alapfogalmak

Vizsgálatunk tárgya az összetett kémiai reakciók mechanizmusa. Ez szemléletesen azt jelenti, hogy állandó térfogatú és nyomású edényt tekintünk állandó hőmérsékleten, amelyben jelen van véges sok (M számú) speciesz (vagy kémiai komponens); legyenek ezek $\mathcal{X}(1), \dots, \mathcal{X}(M)$. A speciesszek között elemi reakciók mennek végbe, amelyet például az alábbiakban, a (2.1) képlet mutat. Azzal fogunk foglalkozni, hogy az elemi reakciók végbemenetele milyen kapcsolatban van az

anyagmegmaradással. Ehhez néhány, meglehetősen absztraktnak tűnő definícióra van szükségünk.

2.1. Mechanizmus, tömegmegőrzés, konzervativitás

2.1. Definíció. [2] (*Összetett kémiai*) mechanizmusnak egy $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ rendezett négyest fogunk nevezni, ahol:

(i) \mathcal{M} és \mathcal{R} egy M , illetve R elemű halmaz ($M, R \in \mathbb{N}$); $\mathcal{R} := \{1, 2, \dots, R\}$; az \mathcal{M} halmaz pedig \mathcal{R} -től diszjunkt, és elemeit $\mathcal{X}(1), \mathcal{X}(2), \dots, \mathcal{X}(M)$ jelöli. (Ahol nem okoz félreértést, ott — de csak ott — az \mathcal{M} halmazt azonosítjuk az $\{1, 2, \dots, M\}$ halmazzal.)

(ii) α és β $M \times R$ -es, nemnegatív egész elemű mátrixok, (elemeik a sztöchiometriai együtthatók), amelyekre teljesül, hogy

- (a) minden $r \in \mathcal{R}$ esetén $\alpha(\cdot, r) \neq \beta(\cdot, r)$,
- (b) ha $\alpha(\cdot, r) = \alpha(\cdot, r')$ és $\beta(\cdot, r) = \beta(\cdot, r')$, akkor $r = r'$,
- (c) minden $m \in \mathcal{M}$ esetén létezik $r \in \mathcal{R}$, amellyel $\alpha(m, r) \neq 0$ vagy $\beta(m, r) \neq 0$.

Az $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmust így szokás ábrázolni:

$$(2.1) \quad \sum_{m=1}^M \alpha(m, r) \mathcal{X}(m) \longrightarrow \sum_{m=1}^M \beta(m, r) \mathcal{X}(m) \quad (r \in \mathcal{R}),$$

és (2.1) r -edik elemére úgy hivatkoznak, mint az r -edik elemi reakcióra. Az \mathcal{M} halmaz $\mathcal{X}(m) \in \mathcal{M}$ elemére úgy fogunk hivatkozni, mint az m -edik *kémiai komponensre* vagy *specieszre*.

A 2.1. definíció (ii) (a) pontjában szereplő feltétel azt jelenti, hogy minden elemi reakciónak van valamilyen hatása. A feltétel egy egyszerű következménye:

$$\forall r \in \mathcal{R} \exists m \in \mathcal{M} \quad \alpha(m, r) \neq 0 \text{ vagy } \beta(m, r) \neq 0.$$

A 2.1. definíció (ii) (b) pontjában szereplő feltétel azt jelenti, hogy a sztöchiometriai együtthatók meghatározzák az elemi reakciót.

A 2.1. definíció (ii) (c) pontjában szereplő feltétel pedig azt jelenti, hogy csak olyan kémiai komponenseket vettünk figyelembe, amelyek legalább egy elemi reakcióban megváltoznak.

A (2.1) mechanizmusban az r -edik *elemi reakciót reverzibilisnek* nevezzük, ha a mechanizmusban szerepel a

$$(2.1') \quad \sum_{m=1}^M \alpha(m, r) \mathcal{X}(m) \longleftarrow \sum_{m=1}^M \beta(m, r) \mathcal{X}(m)$$

elemi reakció is. Egy *reverzibilis mechanizmusban* minden elemi reakció reverzibilis.

Nyilvánvaló, hogy ha $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ egy mechanizmus, akkor $\langle \mathcal{M}, \mathcal{R}, \beta, \alpha \rangle$ is mechanizmus ugyanazokkal a kémiai komponensekkel, és a (2.1')-beli elemi reakciókkal.

Gyakran szükségünk lesz a $\gamma := \beta - \alpha$ mátrixra. Az összetett kémiai mechanizmus definíciója pontosan azt követeli meg γ -tól, hogy ne legyen sem zérus oszlopa, sem zérus sora, elemei egészek legyenek, és ne legyenek azonos oszlopai.

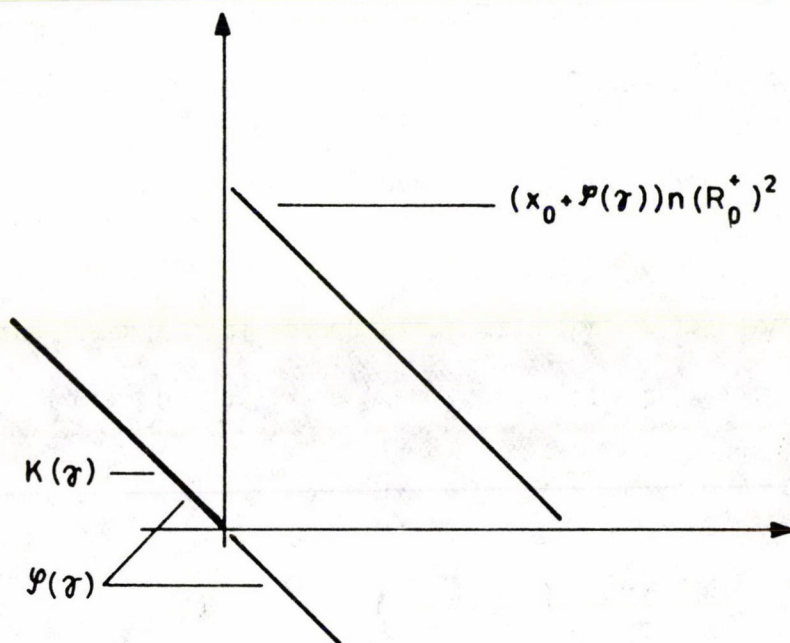
A γ mátrix $\gamma(\cdot, r)$ oszlopai által generált lineáris teret *stöchiometriai térnek* nevezzük, és $S(\gamma)$ -val jelöljük. Tetszőleges $x_0 \in (\mathbb{R}_0^+)^M \setminus \{0\}$ pont esetén az $(x_0 + S(\gamma)) \cap (\mathbb{R}_0^+)^M$ halmaz neve: az x_0 -hoz tartozó reakciószimplex.

Használni fogjuk majd a γ mátrix oszlopai által generált *kúpot* (azaz az oszlopvektorok nemnegatív lineáris kombinációinak halmazát) is:

$$\mathcal{K}(\gamma) := \left\{ \sum_{r \in \mathcal{R}} \lambda_r \gamma(\cdot, r); \quad \forall r \in \mathcal{R} \quad \lambda_r \in \mathbb{R}_0^+ \right\}.$$

Megjegyezzük, hogy $\mathcal{K}(\gamma)$ sohasem csak a 0 vektorból áll a 2.1. definíció (ii) (a) része miatt.

A reakciószimplex és a fenti kúp itteni vizsgátainkban fontos szerepet fog játszani, konkrét dinamikai modellek esetén pedig (amilyeneket itt nem tárgyalunk) a trajektóriák menetéről adnak felvilágosítást.



1. ábra. Reakciószimplex és $\mathcal{K}(\gamma)$ az $X \rightarrow 2Y$ mechanizmusnál

Az r -edik elemi reakció bal és jobb oldalán szereplő formális lineáris kombinációkat (amelyek maguk tehát közvetlenül *nem* elemei egy lineáris térnek, bár megfeleltethetők nekik ilyen elemek) (*reaktáns*, illetve *termék*) *komplexez*eknek nevezik, a megfelelő $\alpha(\cdot, r)$ és $\beta(\cdot, r)$ vektorokat pedig (*reaktáns*, illetve *termék*) *komplex vektoroknak*. Ha egy komplex vektor a nulla vektor, akkor a megfelelő komplexeket *üres komplexez*eknek nevezzük, ennek jele: \emptyset . A komplexeket és a nekik megfelelő komplex vektorokat a továbbiakban azonosítjuk. Ezzel az egyszerűsítéssel élve azt mondjuk, hogy egy η komplex *hossza* (ahol η az $\alpha(\cdot, r)$ és $\beta(\cdot, r)$ vektorok valamelyike:

$\eta \in \{\alpha(\cdot, r); r \in \mathcal{R}\} \cup \{\beta(\cdot, r); r \in \mathcal{R}\}$) koordinátáinak összege: $\ell(\eta) := \sum_{m=1}^M \eta(m)$;

egy komplex *tartója* azon kémiai komponensek indexének halmaza, amelyek az adott komplexben nullától különböző együtthatóval szerepelnek:

$$\text{supp } \eta := \{m \in \mathcal{M}; \eta(m) \in \mathbb{N}\}.$$

2.1. Megjegyzés. Az α és β mátrix elemeit kémiai jelentésük miatt célszerű egészeknek venni. (Valójában sztöchiometriai számításoknál gyakran alkalmaznak törtekkel megadott reakciókat, és ilyenekre vonatkozó sebességi állandókat.) Ezt a dolgot az állításában általában nem fogjuk lényegesen kihasználni, de egyes bizonyításokat valós elemek esetére kissé módosítani kellene.

A kémiai reakció *időbeli lejártszódásának* modelljeiről — azaz a szűkebb értelemben vett reakciókinetikáról — ebben a dolgozatban direkt módon nem esik szó, lásd erről például az [1] vagy a [8] összefoglalót. Amiről itt szó van, az az összes kinetikai modell kerete. Az ilyenfajta vizsgálódások a *sztöchiometria* néven emlegetett területre esnek. A sztöchiometria és kinetika viszonya a mechanikai *kinematika* és *dinamika* viszonyára emlékeztet, amennyiben ott is az előbbi szolgáltatja a keretet az utóbbi leírásához. Még pontosabban: a sztöchiometria foglalkozik a mérleg-egyenletnek, míg a reakciókinetika a transzportegyenletnek megfelelően specializált alakjával.

Dolgozatunk célja az anyag megmaradásának, termelődésének, illetve fogyasztásának vizsgálata. Teljes anyagmérleg készítése esetén a termelődésnek, illetve fogyasztásnak nyilvánvalóan nem lenne értelme; ezek a modellezés közben alkalmazott, de ott igen hasznos fogalmak. Azonkívül azonban, hogy a szokásosan vizsgált formális reakciók nagy részénél (Lotka-Volterra-modell, Oregonátor, explodátor, Brüsszelátor, Schlögl-modell stb.; ezen modellek definíciója és kémiai háttere az eredeti hivatkozásokkal együtt megtalálható például [1]-ben) az anyag formálisan nem marad meg, érdemes még arra is gondolni, hogy az anyag egy része a vizsgált rendszerből távozhathat (például az edény falára), illetve a külvilágból (például az edény faláról) jöhet is be rendszerbe. Másik, igen gyakori ok ilyen „furcsa” modellek létrejöttére az lehet, hogy bizonyos specieszek nagy mennyiségben vannak jelen, s ezek változásától eltekintünk. Lehet, hogy például a *tömegelnyelő* és *-kibocsátó* elnevezés kevésbé lenne fülsértő az ilyen helyzetek megnevezésére.

Az anyagmegmaradást, -termelődést vagy -fogyást két különböző oldalról közelíthetjük meg. Az első esetben a mechanizmus különböző kémiai komponenseit

nem hasonlítjuk össze, hanem csak a reakciók lehetséges lefutásainak következményeit vizsgáljuk. Szemléletesen nyilvánvaló, hogy ha valamely kezdeti állapotból kiindulva a rendszer olyan állapotba jut, hogy egyetlen kémiai komponensből sem tartalmaz kevesebbet, mint a kiinduláskor, de valamely kémiai komponensből többet, akkor a rendszer tömeget termel. A másik megközelítés esetében abból indulunk ki, hogy mérni tudjuk a tömeget. Szemléletesen ismét nyilvánvaló, hogy ha valamennyi reakció egyenlege (vagyis a keletkező anyagok össztömegéből levonva a felhasznált anyagok össztömegét) pozitív (vagyis minden elemi reakció esetében tömeget nyerünk), akkor a mechanizmus képes arra, hogy tömeget termeljen. A két különböző heurisztikus megközelítésből két különböző — egymással duális viszonyban lévő — fogalomrendszerhez jutunk, amelyeket az alábbiakban definiálunk. A következőkben részletesen meg fogjuk vizsgálni a fogalmak kapcsolatait.

2.2. Definíció. Az $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmus

(i) *szigorúan tömegtermelő*, ha

$$(2.2) \quad \mathcal{K}(\gamma) \subset (\mathbb{R}_0^+)^M;$$

(ii) *tömegtermelő*, ha

$$(2.3) \quad \mathcal{K}(\gamma) \cap (\mathbb{R}_0^+)^M \neq \{0\};$$

(iii) *tömegfogyasztó*, ha

$$(2.4) \quad \mathcal{K}(\gamma) \cap (\mathbb{R}_0^-)^M \neq \{0\};$$

(iv) *szigorúan tömegfogyasztó*, ha

$$(2.5) \quad \mathcal{K}(\gamma) \subset (\mathbb{R}_0^-)^M;$$

(v) *tömegmegőrző*, ha se nem tömegtermelő, se nem tömegfogyasztó.

Az utolsó elnevezés csak annyit akar kifejezni, hogy ebben az esetben a mechanizmus se nem termel, se nem fogyaszt anyagot. Az elnevezések mindegyike ideiglenes (és javítandó), ugyanis például az

$$\mathcal{Y} \rightarrow \mathcal{X} \quad 2\mathcal{Y} \rightarrow \mathcal{X}$$

mechanizmus a definíció szerint tömegmegőrzőnek fog minősülni.

2.3. Definíció. Az $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmus

(i) *szigorúan szuperkonzervatív*, ha

$$(2.6) \quad \exists \rho \in (\mathbb{R}^+)^M \quad \rho^T \gamma > 0^T;$$

(ii) *szuperkonzervatív*, ha

$$(2.7) \quad \exists \rho \in (\mathbb{R}^+)^M \quad \rho^T \gamma \vdash \mathbf{0}^T;$$

(iii) *konzervatív* [2], ha

$$(2.8) \quad \exists \rho \in (\mathbb{R}^+)^M \quad \rho^T \gamma = \mathbf{0}^T;$$

(iv) *szubkonzervatív* (vö. [10]), ha

$$(2.9) \quad \exists \rho \in (\mathbb{R}^+)^M \quad \rho^T \gamma \dashv \mathbf{0}^T;$$

(v) *szigorúan szubkonzervatív*, ha

$$(2.10) \quad \exists \rho \in (\mathbb{R}^+)^M \quad \rho^T \gamma < \mathbf{0}^T.$$

2.2. Megjegyzés. A tömeget reprezentáló ρ vektort a 2.3. definícióban nyilván csak szigorúan pozitívnak érdemes venni (ha eltekintünk olyan szélsőséges esetektől, ahol a specieszek között például elektron is szerepel, aminek a tömege esetenként tekinthető nullának atomok vagy molekulák tömegéhez képest). Gyakorlati szempontból a jelző nélküli szub- és szuperkonzervativitás látszik fontosabb fogalomnak; a teljesség és a később kiderülő szimmetria (dualitás) kedvéért vezettük be mind az öt fogalmat.

Megjegyezzük, hogy itt ugyan a definíciók mindegyikében csak γ szerepel, α és β külön nem; a dolgot tárgyan kívül eső témakörökben azonban lényeges szerepe van külön is az α és a β mátrixnak.

Az egyszerűbben kideríthető kapcsolatokra (a fenti két definícióban szereplő két fogalomcsalád belső viszonyaira) már itt rámutatunk, a bonyolultabb eseteket (a két család közötti viszonyokat) a 4. fejezetben tárgyaljuk, a közbülső 3. fejezet pedig szükséges, elégséges, és ekvivalens feltételeket ad meg az egyes tulajdonságok fennállására.

2.3. Megjegyzések. (i) A 2.2. és a 2.3. definícióban *szigorú jelzővel* ellátott tulajdonságok nyilvánvalóan maguk után vonják a jelző nélkülieket. Fordítva ez nem áll, amint azt az Olvasó egyszerű ellenpéldák konstruálásával beláthatja.

(ii) A definíciók egyszerű következménye az is, hogy ha az $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmus teljesíti a 2.2. definícióban szereplő (i) (illetve (ii), (iii), (iv), (v) vagy a 2.3. definícióban szereplő (i), (ii), (iii), (iv), (v)) tulajdonságot, akkor az $\langle \mathcal{M}, \mathcal{R}, \beta, \alpha \rangle$ mechanizmus teljesíti a 2.2. definícióban szereplő (iv), (illetve (iii), (ii), (i), (v) vagy a 2.3. definícióban szereplő (v), (iv), (iii), (ii), (i)) tulajdonságot.

2.4. Megjegyzés. Ha $M = 1$, akkor a szub- és szuperkonzervativitás együtt nem állhat fenn. Ha viszont $M \geq 2$, akkor amellet, hogy egy reakció szigorúan szub- és szigorúan szuperkonzervatív, előfordulhat az is, hogy konzervatív, és az is, hogy nem konzervatív. Ezt mutatják a következők.

2.1. Példák. (i) Tekintsük az $\mathcal{X} \rightarrow \mathcal{Y}$ mechanizmust. Ez a mechanizmus a $\rho_1 := (2, 1)^\top$ vektorral kielégíti a szigorú szub-, a $\rho_2 := (1, 2)^\top$ vektorral kielégíti a szigorú szuperkonzervativitás definícióját, a $\rho_3 := (1, 1)^\top$ vektorral pedig a konzervativitását. A mechanizmus másrészt nyilvánvalóan tömegmegőrző. (ii) Az $\mathcal{X} \rightarrow 2\mathcal{Y}$, $2\mathcal{X} \rightarrow \mathcal{Y}$ mechanizmus pedig a $\rho_1 := (3, 1)^\top$ vektorral kielégíti a szigorú szubkonzervativitást, a $\rho_2 := (1, 3)^\top$ vektorral a szigorú szuperkonzervativitás definícióját, viszont a konzervativitást kifejező egyenletrendszernek nem létezik pozitív megoldása, amint az könnyen látható. Ez a mechanizmus is tömegmegőrző.

2.5. Megjegyzés. A fenti példa által kimutatott anomália a tömegmegőréssel kapcsolatban csak enyhébb formában állhat fenn: egy mechanizmus ugyanis nem lehet egyszerre szigorúan tömegtermelő és szigorúan tömegfogyasztó, de lehet egyszerre tömegtermelő és tömegfogyasztó. Ilyen például a

2.2. Példa. $\mathcal{X} \rightarrow 2\mathcal{X}$, $2\mathcal{X} \rightarrow \mathcal{X}$ mechanizmus (amelyiknek a szokásos rövidítése: $\mathcal{X} \rightleftharpoons 2\mathcal{X}$).

2.6. Megjegyzés. A tömegmegőrésnél (természetesen) az $R = 1$ eset az, amelyik könnyen áttekinthető: ilyenkor a tömegtermelő (tömegfogyasztó) mechanizmus egyúttal szigorúan tömegtermelő (szigorúan tömegfogyasztó) is.

2.2. Atomos szerkezetű specieszeket tartalmazó mechanizmusok

Itt vegyészek számára merésznek (esetleg fölöslegesnek) tűnő absztrakciós lépéseket teszünk. Ezek az absztrakciók viszont egyrészt az alkalmazásokat lefedik, másrészt viszont matematikai szempontból semmilyen további korlátozás nem egyszerűsítene a további vizsgálatokat.

2.3. Példa. A $H_2 + Cl_2 \rightarrow HCl$ mechanizmus mutatja, hogy egy formális mechanizmus lehet úgy is konzervatív, hogy kémiai szempontból teljesen irreális. A kémiailag értelmes $H_2 + Cl_2 \rightarrow 2HCl$ példánál pedig a konzervativitás definíciójában szereplő vektor nemcsak a — molekulasúlyokból álló — $(2, 71, 36.5)^\top$ vektornak vehető, hanem választható például az $(1, 1, 1)^\top$ vektornak is (hiszen ekkor $(-1, -1, 2)^\top$ az egyes kémiai komponensekből keletkezett tömegek vektora, ha a fogyaszt negatív előjellel fejezzük ki), tehát a ρ vektor koordinátáinak nincs feltétlenül kémiai jelentése.

2.4. Definíció. Azt mondjuk, hogy az $(\mathcal{M}, \mathcal{R}, \alpha, \beta)$ mechanizmus *atomos szerkezetű specieszeket tartalmazó mechanizmus*, ha adott egy

(i) D elemű \mathcal{D} halmaz ($D \in \mathbb{N}$), elemei az atomok, $\mathcal{D} = \{\mathcal{A}(1), \dots, \mathcal{A}(D)\}$; és egy

(ii) $D \times M$ -es nemnegatív egész elemű δ mátrix (az *atommátrix*), amelynek egyetlen sor- és oszlopvektora sem a nulla vektor.

Egy atomos szerkezetű specieszeket tartalmazó mechanizmus tehát egy $(\mathcal{M}, \mathcal{R}, \mathcal{D}, \alpha, \beta, \delta)$ rendezett hatos.

A fenti definícióban megfogalmazott viszonyt a komplex mintájára ismét formális lineáris kombinációval lehet kifejezni:

$$(2.11) \quad \mathcal{X}(m) = \sum_{d=1}^D \delta(d, m) \mathcal{A}(d) \quad (m \in \mathcal{M}).$$

A kémiában azonban ezt a kapcsolatot inkább indexekkel szokás kifejezni. Ha például $M := 1$, $\mathcal{X}(1) := H_2O$, akkor $D := 2$, $\delta(1, 1) := 2$, $\delta(2, 1) := 1$, és $\mathcal{A}(1) := H$, $\mathcal{A}(2) := O$; akkor a fenti formális lineáris kombináció így festene; $H_2O = 2H + O$; szokásosan azonban a δ mátrix (egy-től különböző) elemei indexekként jelennek meg.

A fent definiált absztrakt atom az alkalmazásokban lehet például elektron, ion vagy gyök is. Továbbmenve: tetszőleges mechanizmus tekinthető atomos szerkezetű specieszeket tartalmazó mechanizmusnak, ha atomoknak a specieszeket tekintjük, és δ -t az egységmátrixnak vesszük.

A 2.3. példában szereplő első mechanizmust azért neveztük irreálisnak, mert a specieszek atomos szerkezetűek, de az egyes atomok száma az elemi reakcióban megváltozik. Az ilyen átalakulások nem tartoznak a szűkebben vett kémiai vizsgálatok körébe; az itt vizsgálandó elemi reakciókban az atomok száma változatlan. Ennek a tulajdonságnak a formális definíciója következik most.

2.5. Definíció. Legyen $\langle \mathcal{M}, \mathcal{R}, \mathcal{D}, \alpha, \beta, \delta \rangle$ egy atomos szerkezetű kémiai komponenseket tartalmazó mechanizmus. Azt mondjuk, hogy ez a mechanizmus *teljesíti az atomszám-megmaradás törvényét*, ha $\delta\gamma = \mathbf{0}$ ($\in (\mathbb{N}_0)^{D \times R}$).

2.3. Rekeszrendszerek és elsőrendű mechanizmusok

További speciális (nem feltétlenül atomos szerkezetű) mechanizmusokat definiálunk.

2.6. Definíció. Az $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmus *rekeszrendszer* (vagy *kompartment-rendszer*), ha minden $r \in \mathcal{R}$ esetén

$$\sum_{m=1}^M \alpha(m, r) \leq 1 \quad \text{és} \quad \sum_{m=1}^M \beta(m, r) \leq 1.$$

Egy rekeszrendszer *zárt*, ha mindkét helyen egyenlőség áll (vagy, ami ezzel ekvivalens: az α és β mátrix minden oszlopa különbözik a nulla vektortól).

2.7. Definíció. Azt mondjuk, hogy az $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmus *általánosított rekeszrendszer* [6], ha

(i) az $[\alpha \ \beta]$ mátrix minden sorában minden nullától különböző elem értéke azonos, és

(ii) α és β minden oszlopában legfeljebb egy nullától különböző elem áll.

Egy általánosított rekeszrendszer *zárt*, ha az α és β mátrix minden oszlopa különbözik a nulla vektortól.

2.7. Megjegyzés. A definíció más szavakkal azt jelenti, hogy minden kémiai komponens pontosan egy komplexben szerepel, és minden komplex tartója legfeljebb egyelemű, azaz legfeljebb egy kémiai komponenst tartalmaz. Egy általánosított rekeszrendszerben tehát az alábbi típusú elemi reakciók fordulnak elő:

$$(2.12) \quad \eta^m \mathcal{X}(m) \longrightarrow \eta^p \mathcal{X}(p),$$

$$(2.13) \quad \eta^m \mathcal{X}(m) \longrightarrow \mathcal{O}$$

$$(2.14) \quad \mathcal{O} \longrightarrow \eta^m \mathcal{X}(m),$$

ahol $\eta^p, \eta^m \in \mathbb{N}$.

2.8. Megjegyzések. (i) Egy rekeszrendszer egyúttal általánosított rekeszrendszer is, és ha zárt, akkor mindkét értelemben is zárt, ezért sok szempontból elegendő csak általánosított rekeszrendszerekkel foglalkoznunk, bár az alkalmazásokból elsődlegesen a rekeszrendszer fogalma származik, lásd a [9] dolgozatban szereplő hivatkozásokat. (ii) Egy zárt általánosított rekeszrendszerben nyilván csak (2.12) típusú elemi reakciók fordulhatnak elő.

Most a rekeszrendszer fogalmának egy másik általánosítását definiáljuk.

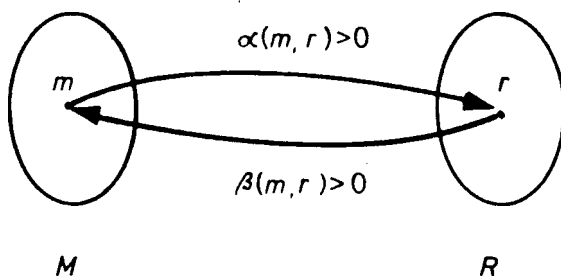
2.8. Definíció. A (2.1)-beli r -edik elemi reakció *rendje* a reaktáns komplex $\ell(\alpha(\cdot, r))$ hossza, egy (összetett kémiai) *mechanizmus rendje* pedig az elemi reakciók rendjének maximuma.

Ennek megfelelően egy összetett kémiai mechanizmust *elsőrendű mechanizmusnak* hívunk, ha a benne szereplő összes elemi reakció rendje legfeljebb 1.

Az elsőrendű mechanizmusok és a rekeszrendszerek viszonyával kapcsolatban megemlíti, hogy minden rekeszrendszer elsőrendű mechanizmus, de fordítva ez nem áll, amint ezt például az $\mathcal{X} \longrightarrow 2\mathcal{Y}$ mechanizmus példája mutatja. Ennek ellenére elterjedt az a szokás, hogy elsőrendű mechanizmust mondanak, és rekeszrendszerre gondolnak.

2.4. A Volpert-gráf

Mechanizmusok vizsgálatánál több gráf is igen hasznosnak bizonyult. Ezek egyikével kapcsolatos fogalmakat definiálunk most.



2. ábra. A Volpert gráf

2.9. Definíció. [10] Az $(\mathcal{M}, \mathcal{R}, \alpha, \beta)$ mechanizmus (redukált) *Volpert-gráfja* az az irányított páros gráf, amelynek csúcshalmaza az $\mathcal{M} \cup \mathcal{R}$ halmaz, és amelynek $\mathcal{X}(m) \in \mathcal{M}$ csúcsából $r \in \mathcal{R}$ csúcsába halad egy él, ha $\alpha(m, r) > 0$, és $r \in \mathcal{R}$ csúcsából $\mathcal{X}(m) \in \mathcal{M}$ csúcsába pedig akkor halad él, ha $\beta(m, r) > 0$. A gráfnak \mathcal{M} -en belül és \mathcal{R} -en belül haladó éle nincs. Egy Volpert-gráfban azon kémiai komponensek $\mathcal{M}_0 \subset \mathcal{M}$ halmazát, amelyekbe nem fut be él, a *gráf kezdőpontjai* halmazának nevezzük.

2.9. Megjegyzések. (i) Ebben a definícióban lényeges, hogy \mathcal{M} és \mathcal{R} diszjunkt, itt tehát \mathcal{M} elemeit nem azonosíthatjuk az első M természetes számmal.

(ii) Aciklikus gráfban minden út kezdőpontja különbözik a végpontjától, ezért nincs benne önmagát metsző út, azaz olyan út, amely valamely ponton legalább kétszer megy át; a gráf végessége következtében léteznek olyan csúcspontok, amelyekbe nem megy be él. Ha egy Volpert-gráf esetén feltesszük, hogy a reaktáns komplex vektorok között a nulla vektor nem szerepel, akkor az ilyen csúcspontok csak kémiai komponenseknek (és nem elemi reakcióknak) megfelelő csúcspontok lehetnek, mivel a feltétel azt jelenti, hogy minden elemi reakciót jelképező csúcsba megy be legalább egy kémiai komponensnek megfelelő csúcsból induló él.

3. A definiált tulajdonságok fennállására vonatkozó szükséges, elégséges, valamint szükséges és elégséges feltételek

3.1. A 2.2. definícióban szereplő fogalmakra vonatkozó feltételek

3.1. Megjegyzés. A 2.2. definícióban szereplő feltételek rendre ekvivalensek a következőkkel:

- (i) $\gamma \vdash 0$;
- (ii) $\forall \mathbf{x}_0 \in (\mathbb{R}_0^+)^M (\mathbf{x}_0 + \mathcal{K}(\gamma)) \cap (\mathbb{R}_0^+)^M$ nem korlátos;
- (iii) $\forall \mathbf{x}_0 \in (\mathbb{R}_0^+)^M (\mathbf{x}_0 - \mathcal{K}(\gamma)) \cap (\mathbb{R}_0^+)^M$ nem korlátos;
- (iv) $\gamma \dashv 0$;
- (v) $\forall \mathbf{x}_0 \in (\mathbb{R}_0^+)^M (\mathbf{x}_0 \pm \mathcal{K}(\gamma)) \cap (\mathbb{R}_0^+)^M$ korlátos.

(ii)-ben, (iii)-ban és (v)-ben elég azt tudni, hogy van olyan $\mathbf{x}_0 \in (\mathbb{R}_0^+)^M$, amelyre a feltétel teljesül.

3.2. Konzervativitásra vonatkozó szükséges és elégséges feltételek

A feltételeket általánosságuk csökkenő sorrendjében adjuk meg.

A sztöchiometriai alteret és a reakciószimplexet felhasználhatjuk a konzervativitás jellemzésére.

3.2. Megjegyzés. Egy mechanizmus akkor és csak akkor konzervatív, ha $S(\gamma)^\perp \cap (\mathbb{R}^+)^M \neq \emptyset$, mivel a konzervatív mechanizmus definíciója egy olyan ρ vektor létezését írja elő, amely mindkét halmaznak eleme.

3.1. TÉTEL [2]. Legyen $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ egy mechanizmus. A következő három állítás ekvivalens:

- (i) A mechanizmus konzervatív.
- (ii) Minden reakciószimplex korlátos.
- (iii) Létezik korlátos reakciószimplex.

Bizonyítás. (i) \Rightarrow (ii) Tegyük fel, hogy a vizsgált mechanizmus konzervatív. Ekkor létezik egy olyan $\rho \in S(\gamma)^\perp$ vektor, amelynek minden koordinátája pozitív. Tekintsünk egy tetszőleges $x_0 \in (\mathbb{R}^+)^M$ ponthoz tartozó reakciószimplexet, és legyen x ennek az $(x_0 + S(\gamma)) \cap (\mathbb{R}_0^+)^M$ reakciószimplexnek egy tetszőleges eleme. Erre teljesül, hogy $\rho^\top (x - x_0) = 0$, ugyanis $x - x_0 \in S(\gamma)$. Ebből következik, hogy $\rho^\top x = K := \rho^\top x_0 > 0$, ugyanis a $\rho^\top x_0$ szorzat első tényezőjének minden koordinátája pozitív, második tényezőjének pedig minden koordinátája nemnegatív, de valamelyik koordinátája nullától különböző. Ha bevezetjük a következő jelölést: $\sigma := \min\{\rho_m; m \in M\}$, akkor — mivel x minden koordinátája nemnegatív — azt kapjuk, hogy $\sigma \sum_{m=1}^M x_m \leq K$. Ez viszont — $\sigma > 0$ miatt — maga után vonja, hogy $0 \leq x_m \leq K/\sigma$ ($m \in M$), azaz a reakciószimplex pontjainak koordinátái közös korlátok közé esnek, a tetszőlegesen kiválasztott reakciószimplex tehát korlátos.

(ii) \Rightarrow (iii) Triviális.

(iii) \Rightarrow (i) Most tegyük fel azt, hogy létezik egy korlátos reakciószimplex; tartozzék ez az x_0 ponthoz. Ekkor $S(\gamma)$ nem tartalmazhat olyan vektort, amelynek minden koordinátája nemnegatív, és különbözik a nulla vektortól. Ha ugyanis x egy ilyen vektor lenne, akkor $x_0 + \lambda x$ tetszőleges $\lambda > 0$ esetén a szimplexben lenne, ezeknek a vektoroknak a normája pedig tetszőlegesen nagy lehet λ alkalmas megválasztásával. Ha tehát a reakciószimplex korlátos, akkor

$$x \in S(\gamma), \quad x \in (\mathbb{R}_0^+)^M \Rightarrow x = 0.$$

Alkalmazva az F.1. tételt és figyelembe véve, hogy a reakciószimplex korlátossága miatt kell, hogy $x = 0$ legyen, azt kapjuk, hogy létezik $\rho \in (\mathbb{R}^+)^M$, $\rho \in S(\gamma)^\perp$, ami éppen azt jelenti, hogy a mechanizmus konzervatív. \square

Természetes módon vetődik föl az a kérdés, hogy adott γ mátrixhoz hány lineárisan független ρ vektor található, amely kielégíti a konzervativitás definíciójában szereplő két feltételt? Nyilvánvalóan vagy egy ilyen vektor sincs, vagy az ilyenek száma éppen $M - \text{rank}(\gamma)$.

3.3. Az aciklikusság következményei

3.2. TÉTEL [10]. Ha egy mechanizmus Volpert-gráfja aciklikus, akkor

(i) ha az α mátrix egyetlen oszlopa sem a nulla vektor (azaz az üres komplex nem szerepel a reaktáns komplexek között), akkor a mechanizmus szigorúan szubkonzervatív;

(ii) ha a β mátrix egyetlen oszlopa sem a nulla vektor (azaz az üres komplex nem szerepel a termék komplexek között), akkor a mechanizmus szigorúan szuperkonzervatív.

Mielőtt bebizonyítjuk a tételt, néhány megjegyzést teszünk.

3.3. Megjegyzések. (i) Ha sem a reaktáns, sem a termék komplexek között nem szerepel az üres komplex, akkor természetesen a reakció egyszerre szigorúan szub- és szuperkonzervatív. Amint korábban már láttuk (2.1. példa), $M \geq 2$ esetén ez előfordulhat.

(ii) Nyilvánvalóan igaz, hogy a szubkonzervativitás és az aciklikusság maga után vonja, hogy az α mátrix egyetlen oszlopa sem a nulla vektor (hiszen ezt már maga a szubkonzervativitás maga után vonja); és természetesen a szubkonzervativitás és az a feltétel, hogy az α mátrix egyetlen oszlopa sem a nulla vektor, *nem* vonja maga után azt, hogy a mechanizmus Volpert-gráfja aciklikus, amint ezt például a $2\mathcal{X} \rightarrow \mathcal{X}$ mechanizmus mutatja. Itt ugyanis a $\rho(-1) < 0$ egyenlőtlenségrendszernek létezik pozitív megoldása: $\rho := 1$; Volpert-gráfja pedig ($\mathcal{X} \leftarrow 1$) ciklikus. (Hasonlókat lehet elmondani a szuperkonzervativitásról és a β mátrixról.)

(iii) Általában is igaz, hogy ha a szigorú szub- vagy szuperkonzervativitás fennáll valamilyen ρ vektorral, akkor létezik $\text{rank}(\gamma)$ számú lineárisan független olyan vektor, amelyekre fennállnak ezek a tulajdonságok, mivel a 2.3. definícióban szereplő szigorú egyenlőtlenségek nyílt halmazokat definiálnak. Így tehát aciklikus Volpert-gráf esetén $\text{rank}(\gamma)$ számú lineárisan független vektorral teljesül a szigorú szub- illetve a szuperkonzervativitás.

A 3.2. tétel bizonyítása. Nyilván elegendő például a tétel első felével foglalkozni, a második fele következik az első feléből, ha figyelembe vesszük a 2.1. megjegyzés (ii) részében mondottakat. Természetesen bebizonyítható a második rész az elsőhöz hasonló módon is.

Mivel a Volpert-gráf aciklikus, kezdőpontjainak halmaza, \mathcal{M}_0 , nem üres.

Tetszőleges $m \in \mathcal{M}$ esetén jelölje $\nu(m)$ a Volpert-gráfban \mathcal{M}_0 -ból m -be vezető leghosszabb út hosszának (ami egy páros szám) a felét; az aciklikusság és az α -ra vonatkozó feltétel következtében $\nu(m)$ értelmezve van minden $m \in \mathcal{M}$ mellett. Vegyük észre, hogy $\alpha(p, r) > 0$ és $\beta(m, r) > 0$ esetén $\nu(p) < \nu(m)$. Rögzítsünk egy C számot, amely nagyobb a β mátrix összes elemének összegénél. Ekkor a

$$\rho_m := C^{-\nu(m)}$$

képlettel definiált ρ vektor garantálja a szigorú szubkonzervativitást, mivel ha rögzített $r \in \mathcal{R}$ -hez választunk olyan $p \in \mathcal{M}$ számot, amelyre $\alpha(p, r) > 0$ teljesül, akkor

$$\begin{aligned} \sum_m \rho_m \beta(m, r) &= \sum_m C^{-\nu(m)} \beta(m, r) \leq C^{-\nu(p)-1} \sum_m \beta(m, r) \\ &< C^{-\nu(p)-1} C = C^{-\nu(p)} = \rho_p \leq \rho_p \alpha(p, r) \leq \sum_m \rho_m \alpha(m, r). \end{aligned} \quad \square$$

Ha α és β elemei nemcsak egész számok lehetnek, akkor ez a bizonyítás kissé módosítandó.

3.4. A rekeszrendszerek különféle általánosításai

3.3. TÉTEL. *Egy általánosított rekeszrendszer pontosan akkor konzervatív, ha zárt.*

Bizonyítás. Tegyük fel, hogy a vizsgált általánosított rekeszrendszer zárt. Ekkor csak (2.12) alakú elemi reakciók fordulhatnak elő benne. Legyen $\rho := (1/\eta^1, \dots, 1/\eta^M)^\top$. Ez a definíció értelmes az általánosított rekeszrendszer definíciójában szereplő (i) feltétel miatt. Az így definiált ρ vektornak nyilván minden koordinátája pozitív, és ha az elemi reakciók közül m -ediknek éppen azt nevezzük, amelyeknek a reaktáns komplexében $X(m)$ szerepel, akkor $\rho^\top \beta = \rho^\top \alpha = 1 \quad (\in \mathbb{R}^R)$, ρ tehát teljesíti a konzervativitás definíciójában szereplő mindkét feltételt.

Ha viszont az általánosított rekeszrendszer nem zárt, akkor van egy olyan elemi reakció (legyen ez az r -edik), amelyikre $\gamma(\cdot, r) = \eta e_p$ valamelyik e_p bázisvektorral és egy nullától különböző η számmal. Ekkor viszont a $\rho^\top \gamma = 0^\top$ egyenletrendszer r -edik egyenlete miatt $\rho_p = 0$ lenne, ellentétben a konzervativitás definíciójával. \square

3.4. TÉTEL. *Tegyük fel, hogy egy elsőrendű mechanizmusban $M = R$ és $\alpha = I_M$. Ezen feltételek mellett a mechanizmus pontosan akkor konzervatív, ha β a kémiai komponensek alkalmas permutálása után az alábbi módon bontható fel:*

$$(3.1) \quad \beta = \begin{pmatrix} P & A \\ 0 & B \end{pmatrix},$$

ahol

- a β mátrix minden oszlopában van zérustól különböző elem,
- P $J \times J$ -s permutáció mátrix (ahol J egy 2 és M közé eső $[2 \leq J \leq M]$ természetes szám),
- 0 zérusmátrix,
- B azon elemei, amelyek nincsenek a főátló fölött, zérusok, egyébként pedig
- A elemei és B többi eleme tetszőleges.

Bizonyítás. A konzervativitás esetünkben azt jelenti, hogy a

$$(3.2) \quad \rho^\top = \rho^\top \beta$$

egyenletrendszernek létezik $\rho \in (\mathbb{R}^+)^M$ megoldása.

1. Egyik irányban tegyük fel, hogy β (3.1) alakú. Ekkor ρ első J koordinátáját 1-nek választva a többi koordináta egymás után kifejezhető a $J + 1$ -edik, ..., M -edik egyenletből. Az így meghatározott ρ vektor K -adik koordinátája (ahol K egy $J + 1$ és M közötti $[J + 1 \leq K \leq M]$ egész szám) a $\rho_1, \dots, \rho_{K-1}$ pozitív számok olyan nemnegatív együtthatós lineáris kombinációja, ahol az együtthatók

között bizonyosan van nullától különböző, így tehát ρ teljesíti a konzervativitásban szereplő mindkét feltételt.

2. Megfordítva pedig azt tegyük fel, hogy a (3.2) egyenletrendszernek létezik pozitív megoldása.

Először megmutatjuk, hogy ha β -nak *nincs* olyan sora, amelynek minden eleme nulla, akkor β permutáció mátrix. Valóban, ilyenkor β minden sorában áll legalább egy pozitív szám. A (3.2) egyenletrendszer egyenleteit összeadva a bal oldalon $\rho_1 + \dots + \rho_M$ -et kapunk, a jobb oldalon pedig minden ρ_m pozitív egész együtthatóval szerepel az összegben, és ez a ρ megoldás pozitív volta miatt csak akkor lehetséges, ha β minden sorában pontosan egy egyes áll. A mátrixban lévő M számú egyes közül viszont kell, hogy minden oszlopba jusson legalább egy (mivel éppen M oszlop van, ezért pontosan egy), hiszen ha például az m -edik oszlop csak nullákat tartalmazna, akkor az m -edik egyenlet nem teljesülne.

Ha pedig β -nak *van* olyan sora, amelyiknek mindegyik eleme 0, akkor számozzuk úgy a kémiai komponenseket, hogy ez legyen az utolsó sor. Ha (3.2) megoldható, akkor az a lineáris egyenletrendszer is megoldható, amit (3.2)-ből az utolsó egyenlet elhagyásával kapunk, ugyanis az első $M-1$ egyenlet — mivel nem tartalmazza ρ_M -et —, nem jelent megkötést ρ_M értékére. Az így kapott egyenletrendszer eggyel kevesebb ismeretlent tartalmaz, ennek együtthatómátrixában megint vagy van olyan sor, amelyiknek minden eleme 0, vagy nincs, s. í. t.

β egyetlen oszlopa sem lehet nulla, mert ez ellentmondana a konzervativitásnak, tehát $J = 0$ nem lehetséges.

Ha $J = 1$ lenne, akkor az elemi reakciók között szerepelne az $\mathcal{X}(1) \rightarrow \mathcal{X}(1)$ elemi reakció, ellentmondásban a 2.1. definíció (ii) (a) részével. \square

Jelöljük most alkalmilag az r -edik elemi reakció reaktáns komplexét így:

$$(\eta^1(r), \dots, \eta^M(r)),$$

termék komplexét pedig így:

$$(\eta^{M+1}(r), \dots, \eta^{2M}(r)).$$

Ekkor a 3.4. tétel feltétele azt jelenti, hogy az elemi reakciók ilyen alakúak:

$$(3.3) \quad \mathcal{X}(m) \rightarrow \mathcal{X}(p) \quad m = 1, \dots, J, \quad m \neq p,$$

(és a jobb oldalon szereplő kémiai komponensek az $\mathcal{X}(1), \dots, \mathcal{X}(J)$ kémiai komponensek egy permutációját adják, miközben végighaladunk az elemi reakciókon);

$$(3.4) \quad \mathcal{X}(m) \rightarrow \sum_{k=1}^{m-1} \eta^{M+m}(k) \mathcal{X}(k), \quad m = J+1, \dots, M.$$

Abban az általánosabb esetben, amikor az $\mathcal{X}(m)$ kémiai komponensből e_m számú elemi reakció indul ki, akkor $e := \max\{e_m; m \in M\}$ számú, (3.2) alakú egyenletrendszer közös pozitív megoldását kell meghatároznunk ahhoz, hogy bebizonyítsuk

az eredeti rendszer konzervativitását. Az egyes egyenletrendszerek egyenleteit a konzervativitást definiáló összes egyenletből az összes lehetséges módon úgy kell kiválasztanunk, hogy minden egyenlet legalább egy rendszerben előforduljon, és hogy mindegyik rendszer pontosan M számú egyenletet tartalmazzon. Ez utóbbi teljesüléséhez az $e_m = 0$ esetben az $x_m = x_m$ egyenletet is a rendszerhez kell csatolnunk.

Az elsőrendű mechanizmus általános esetére tehát egyszerű szükséges és elégséges feltétel helyett egy *eljárást* tudunk megadni, amely e számú lineáris egyenletrendszer együtthatómátrixának a megvizsgálásából áll. Egy egyszerű szükséges és elégséges feltétel, amilyen a fenti 3.4. tételben szerepel, algoritmikusan szintén valamilyen eljárással ellenőrizhető csak (gondoljunk a szénhidrogének pirolízisének előforduló többszáz elemi reakcióra), viszont egy ilyenél az is előfordulhat, hogy a feltételek kémiai nyelven megfogalmazhatók ((3.3) és (3.4) azt jelenti, hogy J számú kémiai komponens egymásba alakul át, a többiekből pedig csak kisebb indexűek keletkezhetnek), s ekkor esetleg kémiai megfontolások alapján eldönthető, hogy teljesülnek-e, vagy sem.

A következő tétel az előzőhöz hasonló módszerrel közvetlenül is bebizonyítható, de az előző tételből is következik, ha β m -edik sorát végigosztjuk az $\alpha(m, m)$ számmal és figyelembe vesszük a 2.1. megjegyzés (i) részét.

3.5. TÉTEL. *Tegyük fel, hogy egy olyan mechanizmusban, amelyben minden reaktáns komplex tartója egy elemű, $M = R$, és az elemi reakciók úgy vannak számozva, hogy α $M \times M$ -es diagonális mátrix. Ezen feltételek mellett a mechanizmus pontosan akkor konzervatív, ha β a kémiai komponensek alkalmas permutálása után az alábbi módon bontható fel:*

$$\beta = \begin{pmatrix} P & A \\ 0 & B \end{pmatrix},$$

ahol

- a β mátrix minden oszlopában van zérustól különböző elem,
- a P $J \times J$ -s mátrix (ahol J egy 2 és M közé eső természetes szám), minden sorában és minden oszlopában pontosan egy nullától különböző elem áll, s ezek az elemek éppen az α mátrix megfelelő sorban álló elemei,
- 0 zérusmátrix,
- A és B pedig olyan, mint a 3.4. tételben.

A jelen szakasz kémiai szempontból legáltalánosabb elégséges feltétele következik, amely matematikai szempontból trivialitás: Ha egy mechanizmus a következő szerkezetű elemi reakciókból áll:

$$(3.5) \quad \eta(m)\mathcal{X}(m) === \sum_{k=1}^{m-1} \eta^{M+m}(k)\mathcal{X}(k), \quad m = J+1, \dots, M;$$

ahol a $===$ jel mindkét irányú nyilat jelenthet, vagy akár mindkettőt egyszerre, és az $\eta(m)\mathcal{X}(m)$ ($m = 1, 2, \dots, J$) komplexek tetszőleges módon vannak egymással elemi reakciókat jelképező nyilakkal összekötve, akkor a mechanizmus konzervatív.

3.5. Az atomszám-megmaradás törvénye

A következőkben megfogalmazzuk azt a matematikailag triviális, kémiai alapvető állítást, amely szerint, ha az atomok a reakciókban megmaradnak, akkor a tömeg is megmarad.

3.6. TÉTEL. *Ha egy atomos szerkezetű kémiai komponenseket tartalmazó mechanizmus teljesíti az atomszám-megmaradás törvényét, akkor az konzervatív.*

Bizonyítás. Legyen $\langle \mathcal{M}, \mathcal{R}, \mathcal{D}, \alpha, \beta, \delta \rangle$ egy atomos szerkezetű kémiai komponenseket tartalmazó mechanizmus. Ha ez a mechanizmus teljesíti az atomszám-megmaradás törvényét, akkor $\delta\gamma = 0$ ($\in \mathbb{R}^{D \times R}$). Legyen $\rho^\top := 1^\top \delta$, ahol $1 \in \mathbb{R}^D$ a csupa egyesekből álló vektor. Az így definiált ρ vektor nyilvánvalóan teljesíti a konzervativitás definíciójában kirótt mindkét feltételt. \square

Amint azt a 2.3. példa is mutatja, a tétel megfordítása nem igaz.

A továbbiakban atomos szerkezetű mechanizmusokkal nem foglalkozunk, lásd ezekről például ARIS, BOWEN, HOLDERITH, PETHŐ, SCHAY és mások [1]-ben felsorolt műveit.

3.6. Kiegészítések

Kimondunk egy általános állítást a reverzibilitás és a 2.2. és 2.3. definícióban szereplő fogalmak viszonyáról.

3.7. TÉTEL. *Ha egy mechanizmus reverzibilis, akkor nem szigorúan szuper-, nem szuper-, nem szub- és nem szigorúan szubkonzervatív; továbbá nem szigorúan tömegtermelő és nem szigorúan tömegfogyasztó.*

Bizonyítás. 1. Az első négy állítás közül nyilván elegendő megmutatnunk, hogy egy reverzibilis mechanizmus nem lehet szuperkonzervatív; a többi állítás elemi logikai műveletekkel, illetve a 2.1. megjegyzés figyelembevételével adódik.

Ha a mechanizmus reverzibilis, akkor minden $r \in \mathcal{R}$ esetén létezik olyan $r' \in \mathcal{R}$, amellyel $\gamma(\cdot, r) = -\gamma(\cdot, r')$. Ennélfogva, ha $\rho^\top \gamma(\cdot, r) > 0^\top$, akkor $\rho^\top \gamma(\cdot, r') < 0^\top$; Vagyis semmilyen tömegvektorral nem lehetnek az elemi reakciókra vonatkozó tömegmérlegek azonos irányúak.

2. A fentiekhez teljesen hasonlóan lehet megmutatni, hogy egy reverzibilis mechanizmus nem lehet szigorúan tömegtermelő sem, mert — amint azt a 4.1. tételben látni fogjuk — akkor szigorúan szuperkonzervatív lenne. \square

Most megadunk egy *algoritmust*, amely eredményesen használható a konzervativitás vizsgálatára.

Tekintsük a következő *lineáris programozási* feladatot. A

$$\rho^\top \gamma = 0^\top, \quad \rho^\top 1 = 1, \quad y1 \leq \rho \quad \rho \geq 0$$

feltételek mellett maximalizálandó az $f(\rho, y) := y$ célfüggvény, ahol ρ \mathbb{R}^M -beli vektor-, y pedig skalárváltozó.

3.8. TÉTEL. Egy mechanizmus pontosan akkor konzervatív, ha a fenti lineáris programozási feladatnak van megengedett megoldása, és az optimális megengedett megoldás(ok)nál az f függvény értéke pozitív.

Bizonyítás. 1. Tegyük fel, hogy a mechanizmus konzervatív. Akkor létezik olyan $\rho > 0$ vektor, amelyre $\rho^T \gamma = 0$ teljesül. Egy ilyen ρ -ból osztással kaphatunk egy olyan ρ^* -ot, amely teljesíti a $\rho^{*T} 1 = 1$ feltételt. Mivel ρ^* koordinátái pozitívak, van alsó korlátjuk, tehát van olyan pozitív y , amellyel $y1 \leq \rho^*$ is teljesül. Van tehát megengedett megoldás, és itt az f függvény értéke pozitív. Az f függvénynek tehát a maximuma is pozitív, hiszen f egy kompakt halmazon definiált folytonos függvény.

2. Ha a fenti lineáris programozási feladatnak van megengedett megoldása, és a célfüggvény maximális értéke pozitív, akkor egy $(\rho^* y^*)$ maximumhelyet véve $\rho^* > y^* 1 > 0$, és $\rho^{*T} \gamma = 0^T$ miatt a mechanizmus konzervatív. \square

3.4. Megjegyzés. Ha egy algoritmus a fenti lineáris programozási feladat optimális megengedett megoldását megengedett megoldásokon keresztül haladva határozza meg, akkor amint a célfüggvény értéke pozitívvá válik, leállítható: a mechanizmus konzervatív.

Tetszőleges mechanizmus egyetlen kémiai komponens hozzávételével konzervatívvá bővíthető:

3.5. Megjegyzés. Tetszőleges $\langle \mathcal{M}, \mathcal{R}, \alpha, \beta \rangle$ mechanizmushoz létezik olyan $\langle \mathcal{M}', \mathcal{R}', \alpha', \beta' \rangle$ konzervatív mechanizmus, amelyre a következők teljesülnek:

$$\begin{aligned}\mathcal{M}' &= \mathcal{M} \cup \{\mathcal{X}(M+1)\}, & \mathcal{R}' &= \mathcal{R}, \\ \alpha'(\cdot, r) &= \alpha(\cdot, r) + (\mathcal{L} - \ell(\alpha(\cdot, r)))e_{M+1}, & r &\in \mathcal{R}; \\ \beta'(\cdot, r) &= \beta(\cdot, r) + (\mathcal{L} - \ell(\beta(\cdot, r)))e_{M+1}, & r &\in \mathcal{R},\end{aligned}$$

itt $\mathcal{L} := \max\{\ell(\alpha(\cdot, r)); r \in \mathcal{R}\} \cup \{\ell(\beta(\cdot, r)); r \in \mathcal{R}\}$. (A konzervativitás definíciójában szereplő ρ vektor $\rho = 1$ -nek vehető.)

3.1. Példa. Az $\mathcal{X}(1) \rightarrow 2\mathcal{X}(1)$ mechanizmus nem konzervatív, de kiegészíthető az $\mathcal{X}(1) + \mathcal{X}(2) \rightarrow 2\mathcal{X}(1)$ konzervatív mechanizmussá. Ennek a kiegészíthetőségnek az a jelentősége, hogy az eredeti mechanizmusra úgy gondolhatunk, mint egy olyanra, amely kémiaiilag nem teljesen irreális, abban az értelemben, hogy legalább konzervatívvá tehető.

4. Kapcsolatok a definiált tulajdonságok között

4.1. TÉTEL. Ha egy mechanizmus

- (i) szigorúan tömegtermelő, akkor szigorúan szuperkonzervatív;
- (ii) szigorúan tömegfogyasztó, akkor szigorúan szubkonzervatív.

Bizonyítás. Nyilván elegendő a tétel első felével foglalkozni, a második fele következik az elsőből, ha figyelembe vesszük a 2.1. (ii) és a 2.3. (ii) megjegyzésben mondottakat. Természetesen bebizonyítható a második rész az elsőhöz hasonló módon is.

A mechanizmus szigorúan tömegtermelő, ezért a 3.1. megjegyzés szerint $\gamma \vdash 0$. Ekkor viszont — mivel a 2.1. definíció (ii) (a) része szerint γ minden oszlopa különbözik a 0 vektortól — *tetszőleges* $\rho > 0$ vektorral teljesül $\rho^\top \gamma > 0^\top$, vagyis a mechanizmus szigorúan szuperkonzervatív. \square

4.2. TÉTEL. Egy mechanizmus

- (i) akkor és csak akkor nem tömegtermelő, ha konzervatív vagy szubkonzervatív;
- (ii) akkor és csak akkor nem tömegfogyasztó, ha konzervatív vagy szuperkonzervatív.

Bizonyítás. Elegendő a tétel első felét bizonyítanunk. Egy mechanizmus pontosan akkor nem tömegtermelő, ha nincs olyan $x \geq 0$, amellyel $\gamma x \vdash 0$. Alkalmazzuk az F.2. tételt a következő szereposztással: legyen $N := R$, és legyen $M_1 := 0$; $M_2 := M$, $B := -\gamma$; $M_3 := R$, $C := -I_R$. Ekkor azt kapjuk, hogy a mechanizmus akkor és csak akkor nem tömegtermelő, ha a

$$B_2\{(y_2, y_3) \in \mathbb{T}; -\gamma^\top y_2 - I_R y_3 = 0, y_2 > 0, y_3 \geq 0\}$$

halmaz nem üres, vagyis ha létezik olyan $y_2 > 0$, $y_3 \geq 0$, amellyel $0 \geq -y_3 = \gamma^\top y_2$. Ha $y_3 = 0$, akkor a mechanizmus konzervatív, míg ha $y_3 \vdash 0$, akkor szubkonzervatív. \square

Az előző tételből közvetlenül adódik, hogy egy mechanizmus akkor és csak akkor tömegmegőrző, ha vagy konzervatív, vagy egyszerre szub- és szuperkonzervatív. Ebből kiemelendő, hogy ha egy mechanizmus konzervatív, akkor tömegmegőrző is egyben. Ez a 3.1. tételből és a 3.1. megjegyzés (v) részéből is következik.

4.3. TÉTEL. Egy mechanizmus pontosan akkor

- (i) szigorúan szubkonzervatív, ha nem tömegtermelő és

$$(4.1) \quad \{x \vdash 0; \gamma x = 0\} = \emptyset,$$

- (ii) szigorúan szuperkonzervatív, ha nem tömegfogyasztó és (4.1) teljesül.

Bizonyítás. Elegendő a tétel első felét bizonyítanunk.

Egy mechanizmus pontosan akkor szigorúan szuperkonzervatív, ha a

$$\{g; \rho > 0, \rho^\top \gamma < 0^\top\}$$

halmaz nem üres. Alkalmazzuk az F.2. tételt a következő szereposztással: legyen $N := M$, és legyen

$$M_1 := M + R, A := \begin{pmatrix} \gamma^\top \\ -I_M \end{pmatrix}; \quad M_2 := 0; \quad M_3 := 0.$$

Ekkor azt kapjuk, hogy a mechanizmus pontosan akkor szigorúan szubkonzervatív, ha a

$$\mathcal{B}_1\{y_1 \in \mathbb{T}; A^\top y_1 = 0, y_1 \vdash 0\} = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{T}; \gamma u - v = 0, \begin{pmatrix} u \\ v \end{pmatrix} \vdash 0 \right\}$$

halmaz üres. Ez a halmaz viszont a következő két részhalmaz egyesítése:

$$\left\{ \begin{pmatrix} u \\ 0 \end{pmatrix} \in \mathbb{T}; \gamma u = 0, u \vdash 0 \right\},$$

(ez szerepel (4.1)-ben), és

$$\left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{T}; \gamma u = v, u \geq 0, v \vdash 0 \right\},$$

Ezen halmaz üres volta ekvivalens azzal, hogy a mechanizmus nem tömegtermelő. \square

A tételből (4.1) nem hagyható el; ezt mutatja a

4.1. Példa. A $\mathcal{X} \longleftrightarrow \mathcal{Y}$ reakció nem tömegtermelő, és nem szigorúan szubkonzervatív.

4.1. Megjegyzés. A (4.1) állítás dinamikai szempontból — például a szokásos determinisztikus modellben értelmezve — azt jelenti, hogy ha egy mechanizmus szigorúan szubkonzervatív, akkor egyetlen állapotból kimozdulva sem juthatunk vissza ugyanoda. Ez tulajdonképpen igen természetes, hiszen ρ -val mérve az anyagokat minden elemi reakció végbemenetele tömegvesztéssel jár. Ha vissza tudnánk térni egy állapotba, akkor valamikor nyerni is kellene tömeget.

4.4. TÉTEL. Egy mechanizmus pontosan akkor

(i) szubkonzervatív, ha nem tömegtermelő és

$$(4.2) \quad \{x > 0; \gamma x = 0\} = \emptyset;$$

(ii) szuperkonzervatív, ha nem tömegfogyasztó és (4.2) teljesül.

Bizonyítás. Elegendő a tétel első felét bizonyítanunk.

Egy mechanizmus pontosan akkor szubkonzervatív, ha

$$\{\rho > 0; \rho^\top \gamma \vdash 0^\top\} \neq \emptyset.$$

Alkalmazzuk az F.2. tételt a következő szereposztással: legyen $N := M$, és legyen $M_1 := M$, $A := -I_M$; $M_2 := R$, $B := \gamma^\top$; $M_3 := 0$. Ekkor azt kapjuk, hogy a mechanizmus pontosan akkor szubkonzervatív, ha a

$$\mathcal{B}_1 = \{(y_1, y_2) \in \mathbb{T}; -y_1 + \gamma y_2 = 0, y_1 \vdash 0, y_2 \geq 0\}$$

és a

$$\mathcal{B}_2 = \{(y_1, y_2) \in \mathbb{T}; -y_1 + \gamma y_2 = 0, y_1 \geq 0, y_2 > 0\}$$

üres. $\mathcal{B}_1 = \emptyset$ azt jelenti, hogy a mechanizmus nem tömegtermelő, $\mathcal{B}_2 \setminus \mathcal{B}_1 = \emptyset$ pedig egyenértékű (4.2) fennállásával. \square

4.2. *Megjegyzés.* A (4.2) állítás dinamikai szempontból — például a szokásos determinisztikus modellben értelmezve — azt jelenti, hogy ha egy mechanizmus szubkonzervatív, akkor ha *minden* elemi reakció legalább egyszer végbemegy, akkor ennek során egyetlen állapotból kiindulva sem juthatunk vissza az eredeti állapotba. Ez tulajdonképpen igen természetes, hiszen valamelyik elemi reakció végbemenetele határozott tömegvesztéssel járt, és egyetlen elemi reakcióban sem nyertünk tömeget.

5. Kitekintés

Megemlíttünk néhány, a továbbiakban megoldandó részproblémát.

1. Hogyan ellemőrizhetők a 3.4. és a 3.5. tétel feltételei, vagyis hogyan állapítható meg az α és a β mátrixról, hogy alkalmas permutálás után megfelelő alakra hozható-e. A tételek bizonyításából is kiolvasható egy algoritmus, de ilyen kérdésekkel foglalkozik SOLTÉSZ P. [4] szakdolgozata is.

2. *Számelméleti* terminológiával élve a konzervativitás azt jelenti, hogy létezik a

$$\rho^T \alpha = \rho^T \beta$$

diofantoszi egyenletrendszernek pozitív egész koordinátákkal bíró ρ megoldása. (Nyilván pontosan ugyanakkor, amikor pozitív valós megoldás létezik, ha α és β elemei egészek.) Standard tankönyvek nem adnak erre nézve használható felvilágosítást, talán ez az átfogalmazás mégis sugallhat valamit.

3. Kémiai szempontból a legfontosabb az lenne, hogy másodrendű mechanizmusokra kapjunk minél explicitebb feltételeket. Esetleg feltehető, hogy a termék-komplexek hossza sem hosszabb kettőnél; ugyanis ez az eset is elég érdekes lenne.

4. A 3. fejezet állításainak megfelelői feltehetően megfogalmazhatók a tömeg-megőrzéssel kapcsolatos fogalmakra is.

FÜGGELÉK

TUCKER alább idézett állítása (kicsit más formában) megtalálható például a [3] könyvben.

F.1. TÉTEL. Legyen $M \in \mathbb{N}$, és legyen S az \mathbb{R}^M vektortér altere. ekkor létezik olyan $x \in (\mathbb{R}_0^+)^M$, $\rho \in (\mathbb{R}_0^+)^M$, amelyekre a következők teljesülnek:

$$x \in S, \quad \rho \in S^\perp, \quad x + \rho \in (\mathbb{R}^+)^M. \quad \square$$

Idézzük SZILÁGYI [5] egy teljesen általános alternatíva tételének számunkra elegendő speciális esetét.

F.2. TÉTEL. Legyen $N \in \mathbb{N}$, $M_1, M_2, M_3 \in \mathbb{N}_0$ és legyen

$$A \in \mathbb{R}^{M_1 \times N}, B \in \mathbb{R}^{M_2 \times N}, C \in \mathbb{R}^{M_3 \times N}; \quad T := \mathbb{R}^{M_1} \times \mathbb{R}^{M_2} \times \mathbb{R}^{M_3}.$$

Ekkor az alábbi A és B halmaz közül pontosan egy üres:

$$A := \{x \in \mathbb{R}^N; Ax < 0, Bx \neq 0, Cx \leq 0\}, \quad B := B_1 \cup B_2,$$

ahol

$$B_1 := \{(y_1, y_2, y_3) \in T; A^T y_1 + B^T y_2 + C^T y_3 = 0, y_1 \neq 0, y_2 \geq 0, y_3 \geq 0\},$$

$$B_2 := \{(y_1, y_2, y_3) \in T; A^T y_1 + B^T y_2 + C^T y_3 = 0, y_1 \geq 0, y_2 > 0, y_3 \geq 0\}. \quad \square$$

F.1. Megjegyzések. (i) A fentieket pontosabban úgy értjük, hogy ha $M_1=0$, akkor B_1 , ha $M_2=0$, akkor B_2 nem szerepel a tétel állításában. (ii) A B_1 és B_2 halmaz definíciója csak finom részletekben tér el egymástól. Számos közös elemük lehet: mindazok az $(y_1, y_2, y_3) \in T$ hármasok, amelyek az egyéb feltételeket kielégítik és amelyekre $y_1 \neq 0$ és $y_2 > 0$.

Köszönetnyilvánítás. A kézirat korábbi változatainak alapos áttanulmányozása után számos értékes javaslatot adott Farkas Henrik és Póta György. A dolgozat részben a 3268. számú OTKA-pályázat támogatásával készült.

IRODALOM

- [1] ÉRDI, P. és TÓTH, J., *Mathematical models of chemical reactions* (Manchester University Press, Manchester, és Princeton University Press, Princeton, 1989).
- [2] HORN, F., és JACKSON, R., „General mass action kinetics”, *Arch. Ratl. Mech. Anal.* **47** (1972), 81–116.
- [3] PRÉKOPA A., *Lineáris programozás* (Bolyai János Matematikai Társulat, Budapest, 1968).
- [4] SOLTÉSZ P., *Szakkolgozat* (ELTE TTK, Budapest).
- [5] SZILÁGYI P., *Consistency of linear systems*, In: Fourth Conf. of Program Designers, (A. Iványi), (Budapest, 1988), 147–152.
- [6] TÓTH J., „A kémiai reakciókinetika direkt és inverz feladatairól”, *Alk. Mat. Lapok* **7** (1981), 253–269.
- [7] TÓTH J., „A formális reakciókinetika globális determinisztikus és sztochasztikus modelljéről és néhány alkalmazásáról”, *MTA SZTAKI Tanulmányok* **129** (1981), 1–163.
- [8] TÓTH J. és ÉRDI P., „A formális reakciókinetika modelljei, problémái és alkalmazásai”, *A kémia újabb eredményei* **41** (1978), 226–352.
- [9] TÓTH J. és HÁRS V., „A rekeszrendszerek inverz feladatáról”, *Alk. Mat. Lapok* **5** (1979), 49–61.
- [10] VOL'PERT, A. I., HUDJAEV S. I., *Analysis in Classes of Discontinuous Functions and Equations of Mathematical Physics* (Martinus Nijhoff Publishers, Dordrecht – Boston – Lancaster, 1985), (Oroszul: Nauka, Moszkva, 1975).

(Beérkezett: 1990. május 18.)

DEÁK JENŐ
MTA MATEMATIKAI KUTATÓ INTÉZET
BUDAPEST, PF. 127, 1364

TÓTH JÁNOS
GATE SZÁMÍTÁSTECHNIKAI TANSZÉK
GÖDÖLLŐ, PÁTER K. U. 1., 2103
TOTHJ@SZKP.GAU.HU
VIZVÁRI BÉLA
ELTE TTK OPERÁCIÓKUTATÁSI TANSZÉK
1088 BUDAPEST, MÚZEUM KRT. 6-8.
VIZVARI@CS.ELTE.HU ÉS
BILKENT UNIVERSITY, 06533
ANKARA
VIZVARI@TRBILUN.BITNET

CONSERVATION OF MATTER IN COMPLEX CHEMICAL MECHANISMS

J. DEÁK, J. TÓTH and B. VIZVÁRI

This paper aims at putting forward a unified theory of mass conservation. In addition to mass conservativity and related properties of complex chemical mechanisms, we introduce a new system of concepts, which can be regarded as dual to the old one. After reviewing previous results (with proofs and examples) on mass conservativity, we give necessary and/or sufficient conditions that are simpler or easier to handle, as well as algorithms for checking mass conservativity. Based on a general alternative theorem for linear inequalities, we prove results connecting the two systems of concepts.

A SZÜKSÉGES ÉS ELÉGSÉGES ÖSSZEFÉRHETŐSÉGI PEREMFELTÉTELEK MEGHATÁROZÁSA

LÁMER GÉZA

Budapest

A tanulmányban a *Southwell-paradoxon* megoldása során felmerült összeférhetőségi peremfeltételekkel foglalkozunk. Megadjuk a szükséges és elégséges összeférhetőségi peremfeltételeket. Bebizonyítjuk, hogy a WASHIZU által adott megoldás nem csak szükséges, hanem elégséges is.

1. Bevezetés

A háromdimenziós kontinuummechanikában bizonyítást nyert, hogy az alakváltozási tenzor hat (független) komponenséből az (egy merevtestszerű elmozdulástól eltekintve) egyértékű eltolódásvektor előállításának feltétele a hat összeférhetőségi mezőegyenlet teljesülése (lásd pl. LUR'E [5] 60-64. old. a linearizált, és 87-89. old. a nemlineáris esetre). Az egyértékűség mindig a zárójelbe tett megjegyzéssel együtt értendő, így ennek kiírásától a továbbiakban eltekintünk.

Itt, és a továbbiakban is, feltesszük, hogy a vizsgált V tartomány egyszeresen összefüggő, az S pereme pedig reguláris, korlátos és zárt.

Az egyértékű eltolódásmezőt nemcsak a V tartományon, hanem annak S peremén is elő kell tudni állítani, hiszen a peremérték-feladat a $\bar{V} = V \cup S$ zárt tartományra vonatkozik. Az inkompatibilitási tenzor komponenseinek az eltűnését az S peremen összeférhetőségi peremfeltételeknek nevezzük. Az inkompatibilitási tenzor komponenseinek folytonosságából következik, hogy ha az összes összeférhetőségi mezőegyenlet teljesül, akkor az összes összeférhetőségi peremfeltétel is teljesül. Ezért az összeférhetőségi peremfeltételek automatikusan teljesülnek.

A kontinuummechanika duál feladatában ettől eltérő helyzet áll elő. Ugyanis a hat összeférhetőségi mezőegyenlet helyett csak hármat alkalmazunk, így további feltételeket kell előírni ahhoz, hogy az összeférhetőségi feltételek teljesüljenek mind a V tartományon, mind annak S peremén.

A kontinuummechanika duál feladatában tehát a következő három probléma merül fel.

1°. A hat elsődleges ismeretlen meghatározásához a hat összeférhetőségi mezőegyenletből melyik hármat kell az egyensúlyi (mozgás-) egyenletekhez csatolni.

2°. Milyen feltételnek kell fennállnia ahhoz, hogy végül is mind a hat összeférhetőségi feltétel — mint az eltolódásmező egyértékűségének feltétele — a $\bar{V} = V \cup S$ zárt tartományon teljesüljön.

3°. A primál rendszerben is fennálló pontonkénti három peremfeltételhez milyen további három peremfeltételt kell csatolni ahhoz, hogy a duál feladat feszültségmezővel felírt peremérték-feladatában a szükséges hat peremfeltétel adott legyen.

A lineáris rugalmasságtanban a variációs elvek és a feszültségfüggvények együttes alkalmazása a fenti problémákkal analóg problémákra vezetett; az úgynevezett *Southwell-paradoxonra*. Ezért először ismertetjük a *Southwell-paradoxont*, és a feloldását.

2. A Southwell-paradoxon és feloldása

Az ebben a paragrafusban ismertetett paradoxon és feloldásai a lineáris rugalmasságtan összefüggéseire vonatkoznak.

A kontinuummechanikában bevezett másodrendű szimmetrikus inkompatibilitási tenzort ϱ -val fogjuk jelölni.

2.1. SOUTHWELL [6] statikai feladatokat vizsgálva megmutatta, hogy ha a *Castigliano-elvet* a *Mazwell-féle feszültségfüggvénnyel* kombinálja, akkor mezőegyenletként a

$$(2.1) \quad \varrho^{11} = \varrho^{22} = \varrho^{33} = 0, \quad \text{a } V\text{-n}$$

míg ha a Morera-féle feszültségfüggvénnyel kombinálja, akkor a

$$(2.2) \quad \varrho^{12} = \varrho^{23} = \varrho^{31} = 0, \quad \text{a } V\text{-n}$$

összeférhetőségi mezőegyenleteket nyeri.

A „paradoxon” lényege a következő: Az alakváltozásból az eltolódásvektor egyértelmű meghatározásának a szükséges és elégséges feltétele a hat összeférhetőségi mezőegyenlet teljesülése. A feszültségeknek a három feszültségfüggvénnyel történő megadása is szükséges és elégséges feltétel. (Pontosabban, úgy fogalmaz, hogy a *Mazwell*-, illetve *Morera-féle feszültségfüggvények* a legáltalánosabb alakú feszültségfüggvények, amelyek az egyensúlyi egyenleteket kielégítik.) A feszültségfüggvénnyel felírt *Castigliano-elv* csak három összeférhetőségi mezőegyenletre vezet. Tehát az eltolódásmező egyértelműségének biztosításához az alakváltozás-eltolódás összefüggések integrálásából hat, míg variációs elvből, feszültségfüggvények alkalmazása esetén, három szükséges és elégséges összeférhetőségi mezőegyenlet adódik (SOUTHWELL [7]).

2.2. WASHIZU [11] bebizonyította, hogy ha feltesszük, hogy

$$(2.3) \quad \varrho^{11} = \varrho^{22} = \varrho^{33} = 0, \quad \text{a } V\text{-n}$$

$$(2.4) \quad \varrho^{12} = \varrho^{23} = \varrho^{31} = 0, \quad \text{az } S\text{-en}$$

teljesül, akkor a

$$(2.5) \quad \varrho^{12} = \varrho^{23} = \varrho^{31} = 0, \quad \text{a } V\text{-n}$$

fennállása már következik. Bebizonyította azt is, hogy ha feltesszük, hogy

$$(2.6) \quad \varrho^{12} = \varrho^{23} = \varrho^{31} = 0, \quad \text{a } V\text{-n}$$

$$(2.7) \quad \varrho^{11} = \varrho^{22} = \varrho^{33} = 0, \quad \text{az } S\text{-en}$$

teljesül, akkor a

$$(2.8) \quad \varrho^{11} = \varrho^{22} = \varrho^{33} = 0, \quad \text{a } V\text{-n}$$

fennállása már következik. Rámutatott arra is, hogy ha feltesszük, hogy

$$(2.9) \quad \varrho = 0, \quad \text{az } S\text{-en}$$

teljesül, akkor bizonyítható, hogy az összeférhetőségi mezőegyenletekből hármat, nevezetesen a (2.3), vagy a (2.6) egyenleteket elegendő a hat összeférhetőségi feltétel közül a mezőegyenletekhez csatolni.

WASHIZU a *Southwell-paradoxont* feloldotta: A *Castigliano-elvből* lehetséges hat helyett három összeférhetőségi feltételt nyerni mint mezőegyenletet, a további (fennmaradó) három összeférhetőségi feltételt a peremre kell „áthelyezni”.

WASHIZU a *Southwell-paradoxon* feloldása kapcsán felhívja a figyelmet a következő problémára: Mi a szükséges és elégséges összeférhetőségi feltétel a peremen, amelyet a (2.3), vagy a (2.6) egyenlethez csatolni kell, mivel a (2.9) „would seem to be redundant”, azaz fölöslegesnek látszik. WASHIZU szerint a problémát a variációs eljárás során a peremfeltételek tüzetes vizsgálatával lehet megoldani.

2.3. KOZÁK [3] a *Southwell-paradoxont* megoldottnak tekinti. Mégis az összeférhetőségi peremfeltételeket WASHIZUétól eltérő

$$(2.10) \quad \varrho n = 0, \quad \text{az } S\text{-en}$$

alakban veszi fel. Itt n az S felület külső normális egységvektora. Ezt követően bebizonyítja, hogy ha feltesszük, hogy

$$(2.11) \quad \varrho^{XY} = 0, \quad \text{a } V\text{-n}$$

$$(2.10) \quad \varrho n = 0, \quad \text{az } S\text{-en}$$

teljesül, akkor a

$$(2.12) \quad \varrho^{AB} = 0, \quad \text{a } V\text{-n}$$

fennállása már következik. Itt az XY és AB indexpárok kiadják a szimmetrikus inkompatibilitási tenzornak mind a hat független koordinátáját. Egyúttal megadja az XY és AB indexpárok kiválasztási szabályát, amely azonos azzal a szabállyal, ahogyan a szimmetrikus feszültségfüggvény-tenzor hat független koordinátájából kell hármat kiválasztani a feszültségek előállításához (ez utóbbit lásd BLOH [1] 312–315. old.). A (2.10) összefüggést, megkülönböztetésül az összeférhetőségi peremfeltételtől, *Kozák-féle összeférhetőségi peremfeltételnek* fogjuk nevezni.

KOZÁK [3] dolgozatában a *Southwell-paradoxon* feloldását WASHIZU megoldásától eltérő, de annál általánosabb formában adta meg, mivel a bizonyítást tetszőleges görbevonali koordinátarendszerben és az inkompatibilitási tenzor három tetszőleges (a kiválasztási szabálynak eleget tevő) komponensére végezte el. Ezzel együtt általánosította WASHIZU a *Descartes-féle koordinátarendszerben* megadott, a *Maxwell- és Morera-féle feszültségfüggvényekhez* kapcsolódó megállapítását a feszültségfüggvényeknek mind az öt családjára (lásd BLOH [1] 314. old.), valamint tetszőleges görbevonali koordinátarendszerben felírható feszültségfüggvények eseteire is.

Ugyanakkor a KOZÁK által adott megoldás több ponton is eltér a WASHIZU által adotttól. Egyrészt a peremfeltételt a perem függvényében adta meg. Másrészt a (2.10) nem ekvivalens sem a WASHIZU által alkalmazott (2.4), sem pedig a (2.7) feltétellel; de a legáltalánosabb alakú (2.9) feltétellel sem. Ezért KOZÁK megoldása — ebből a szempontból — független WASHIZU megoldásától. Figyelembe véve WASHIZU megjegyzését, meg kell vizsgálni, vajon a *Kozák-féle összeférhetőségi feltétel* szükséges-e.

2.4. Áttekintve WASHIZUNAK és KOZÁKNAK a *Southwell-paradoxonra*, azaz a bevezetésben felsorolt három probléma megoldására vonatkozó eredményeit, megállapítható, hogy

WASHIZU megoldása

- speciális esetekre megfogalmazott (kétféle kiválasztási szabály, *Descartes-féle koordinátarendszer*),
- független a V tartományt határoló S peremtől,
- az inkompatibilitási tenzornak csak három komponense szerepel az összeférhetőségi peremfeltételekben,
- nincs tisztázva, hogy ez az összeférhetőségi peremfeltétel elégséges-e;

KOZÁK megoldása

- általánosságban megfogalmazott (általános kiválasztási szabály, tetszőleges görbevonali koordinátarendszer),
- függ a V tartományt határoló S peremtől (az n normálvektoron keresztül),
- az inkompatibilitási tenzornak mind a hat komponense szerepel az összeférhetőségi peremfeltételben,
- nincs tisztázva, hogy ez az összeférhetőségi peremfeltétel szükséges-e.

Összehasonlítva WASHIZU és KOZÁK megoldásait, megállapítható, hogy

- mivel (2.9)-ből következik (2.10), ezért KOZÁK igazolta WASHIZU bizonyítás nélkül közölt állítását (a fenti értelemben vett általános esetre is).

– mivel (2.10)-ből nem következik sem (2.4), sem (2.7), ezért KOZÁK megoldása a két, WASHIZU által tárgyalt speciális esetben nem adja vissza WASHIZU megoldását. (Ebből a szempontból a két megoldás független egymástól.)

2.5. A *Southwell-paradoxon* feloldása kapcsán tehát két probléma maradt nyitottan:

1°. Mi is a felületen megadott alakváltozásból egyértékű eltolódásmező meghatározásának a szükséges és elégséges feltétele? Másként megfogalmazva, mi a szükséges és elégséges összeférhetőségi feltétel egy felületen?

2°. Mi is a kontinuummechanika duál feladatában a szükséges és elégséges összeférhetőségi peremfeltétel?

A következő paragrafusban bebizonyítjuk, hogy a szükséges és elégséges összeférhetőségi feltétel egy felületen az inkompatibilitási tenzor eltűnése a felületen.

A 4. paragrafusban bebizonyítjuk, hogy a hat összeférhetőségi feltétel helyettesítése (megfelelően kiválasztott) három összeférhetőségi mezőegyenlettel és három összeférhetőségi peremfeltétellel

– szükséges és elégséges ahhoz, hogy mind a hat összeférhetőségi mezőegyenlet teljesüljön,

– szükséges és elégséges ahhoz, hogy mind a hat összeférhetőségi peremfeltétel teljesüljön.

A javasolt megoldásból az is következik, hogy a három (előírt) összeférhetőségi peremfeltétel elégséges a duál módszer peremfeltétel-rendszerének teljessé tételéhez.

Megjegyezzük, hogy a „megfelelő kiválasztás” módját KOZÁK [3] megmutatta.

3. Az összeférhetőségi feltételekről

3.1. Összeférhetőségi mezőegyenletek

A kontinuummechanika szempontjából, mint már említettük, az egyértékű eltolódásvektor előállításának szükséges és elégséges feltétele egy egyszerűen összefüggő tartományon a hat összeférhetőségi egyenlet teljesülése, azaz az inkompatibilitási tenzor hat (független) komponensének az eltűnése a vizsgált tartományon.

Differenciálgeometriai szempontból az egyértékű eltolódásvektor előállításának szükséges és elégséges feltétele a kovariáns (abszolút) deriváltak felcserélhetősége, azaz a *Riemann-Christoffel-féle negyedrendű görbületi tenzor* eltűnése a vizsgált tartományon (lásd LUR'E [5] 87–89. old.). Ekkor a vizsgált tér (tartomány) euklideszi (egyenes) (lásd pl. KORN–KORN [2] 526–527. old., LÁNCZOS [4] 160–162. és 171–173. old.). A *Riemann-Christoffel-féle görbületi tenzornak* a deformálatlan és deformált állapotbeli eltűnéséből következik, hogy a *Riemann-Christoffel-féle görbületi tenzornak* a deformáció során létrejött megváltozása — azaz a negyedrendű inkompatibilitási tenzor — zérus értékű. A differenciálgeometria szempontjából ezek lesznek az összeférhetőségi egyenletek.

A negyedrendű inkompatibilitási tenzornak a háromdimenziós esetben hat független komponense van (analog a *Riemann-Christoffel-féle görbületi tenzorhoz*, ez

utóbbit lásd pl. KORN-KORN [2] 525. old., vagy LÁNCZOS [4] 167–168. old.). Ezekből, megfelelően megválasztott indexezés mellett, az első és a második indexpárban a $23 \Rightarrow 1$, $31 \Rightarrow 2$ és $12 \Rightarrow 3$ jelölést bevezetve, egy szimmetrikus, másodrendű tenzor alkotható (analóg a *Ricci-féle tenzorhoz*, lásd pl. LUR'E [5] 889–890. old.). Ez a másodrendű inkompatibilitási tenzor azonos a kontinuummechanikában bevezetett inkompatibilitási tenzorral.

Az összeférhetőségi mezőegyenleteket a V tartományon tehát a

$$(3.1) \quad \varrho = 0, \quad \text{a } V\text{-n}$$

alakban írhatjuk fel.

Rá kell mutatni arra, hogy a kontinuummechanika szorosan kapcsolódik az euklideszi tér fogalmához: a vizsgált test térben létrejövő helyváltozását az eltolódásvektorral jellemezzük. Az eltolódásvektor csak az euklideszi térben vezethető be, (nemzérus görbületű) *Riemann-térben* nem. (Az euklideszi tér egyúttal lineáris vektortér is, míg a *Riemann-tér* nem.) Ezért szükséges, hogy mind a *Riemann-Christoffel-féle görbületi tenzor*, mind pedig annak megváltozása zérus legyen.

Megjegyzés. Az inkompatibilitási tenzornak a differenciálgeometriai megfogalmazása lehetővé teszi az összeférhetőségi egyenletek alkalmazását az alakváltozás-eltolódás összefüggések tetszőleges nemlineáris alakjára is.

A következő paragrafusban szükségünk lesz a *Bianchi-azonosságra*: a *Riemann-Christoffel-féle görbületi tenzor* (az $R = 0$ összefüggéstől függetlenül) kielégíti a *Bianchi-azonosságot* (lásd pl. KORN-KORN [2] 525. old., vagy LÁNCZOS [4] 167–168. old.). Ez a megállapítás a deformációt megelőző, illetve követő állapotban is fennáll, ezért az inkompatibilitási tenzor is kielégíti a *Bianchi-azonosságot*. A *Bianchi-azonosságot* a másodrendű inkompatibilitási tenzorra (a V tartományon) a

$$(3.2) \quad \nabla_i \varrho^{ij} = 0, \quad \text{a } V\text{-n}$$

alakban írjuk fel. Itt $i = j = 1, 2, 3$ és alkalmazzuk az *Einstein-féle összegzési megállapodást*.

3.2. Összeférhetőségi feltételek egy felületen

Mielőtt meghatároznánk az összeférhetőségi feltételeket egy felületen a következő két lehetőséget kell elkülöníteni egymástól.

1°. Csak a felületet tekintjük. Csak a felületen adott az alakváltozási mező, azaz csak a felület deformálódott. Tehát a csak a felületen adott alakváltozási mezőből határozzuk meg a felület és csak a felület eltolódását.

2°. A háromdimenziós tér egy háromdimenziós zárt \bar{V} tartománya deformálódását vizsgáljuk, de ezen belül kijelölünk egy S felületet. Tehát az alakváltozási mező a vizsgált háromdimenziós tartományon adott, de egyenlőre csak a kijelölt felületen kívánjuk az eltolódásmezőt előállítani, természetesen a felületen megadott

alakváltozási mezőből. (Ez utóbbi természetesen része a felületet magába foglaló tartományon megadott alakváltozási mezőnek.)

Az első esetet ki kell zárni. Ugyanis a kontinuummechanikában az alakváltozás a metrikus tenzorral van kapcsolatban (lásd pl. LUR'E [5] 73–75. old.), a felületelméletből viszont ismert, hogy a felület egyértelmű meghatározásához (eltekintve a merevtestszerű elmozdulásoktól) nem elegendő a felület (pozitív definit) metrikus tenzorát (azaz az első alapformát) ismerni, hanem a görbületi tenzor (azaz a második alapforma) ismerete is szükséges (lásd a felületelmélet alaptételét pl. SULIKOVSKIJ [9] 262. old.).

A második esetben a \bar{V} tartományon az eltolódásmezőt két lépésben állítjuk elő. Először az S felületen, majd a $\bar{V} \setminus S$ tartományon. A kétféleképp számított eltolódásmezőknek a teljes \bar{V} tartományon kell kompatibilisnek lenniük. A teljes \bar{V} tartományon a szükséges és elégséges feltétel természetesen a (3.1) alatti hat összeférhetőségi mezőegyenlet teljesülése. Ebből következik, hogy az összeférhetőségi feltételek teljesülése az S felületen elégséges feltétel lesz a teljes eltolódásmező kompatibilitásához. Kérdés, hogy vajon szükséges is ez? Az, hogy az S felületen az inkompatibilitási tenzor eltűnése, mint feltétel, szükséges is, az az alábbiakból következik: a vizsgált háromdimenziós tartomány euklideszi térben adott, tehát megadható a tartomány egészét egyben lefedő megengedett koordinátarendszer. Ezt tekinthetjük úgy is, hogy egy felületsereg (egyrétegűen) fedi le a tartományt. Ha tehát egy felületen a hat összeférhetőségi feltételnél kevesebb is elegendő volna az eltolódásmező egyértékű előállításához, akkor az euklideszi térben a tartományt lefedő (egyrétegű) felületseregen is elegendő lenne a hat összeférhetőségi feltételnél kevesebb is az eltolódásmező egyértékű előállításához. Az ellentmondás miatt a felületen a hat összeférhetőségi feltétel teljesülése szükséges is. Azaz egy reguláris felületen megadott alakváltozási mezőből egyértékű eltolódásmező előállításának szükséges és elégséges feltétele az inkompatibilitási tenzor komponenseinek az eltűnése a felületen. Ezt a

$$(3.3) \quad \varrho = 0, \quad \text{az } S\text{-en}$$

alakban írjuk fel.

A kontinuummechanikában az $S = \partial V$ választással nyerjük az összeférhetőségi peremfeltételeket.

Megjegyzés. Az alakváltozási mező és az eltolódásvektor értelmezési tartománya dimenziószámának csökkentése a probléma lényegét nem érinti: mind az alakváltozási mező, mind pedig az előállítandó eltolódásvektor-mező a háromdimenziós beágyazó (euklideszi) térben van értelmezve, ebben a térben kívánunk egyértelmű eltolódásmezőt előállítani, tehát ennek a térnek kell zérus görbületűnek lennie.

3.3. Appendix

KOZÁK [3] a linearizált elméletben egy reguláris felületen megadott alakváltozási mezőből az egyértékű eltolódásmező előállításának szükséges és elégséges feltételét az adott felület geometriájához kötődő

$$(3.4) \quad \varrho n = 0, \quad \text{az } S\text{-en}$$

alakban állította elő.

A fenti fejtegetés értelmében ez akkor és csak akkor lehet szükséges és elégséges feltétel, ha (3.3) és (3.4) következnek egymásból.

$A \varrho = 0 \Rightarrow \varrho n = 0$ implikáció triviális.

A $\varrho n = 0 \Rightarrow \varrho = 0$ implikáció tetszőleges szimmetrikus tenzort tekintve nem áll fenn. (Három egyenlet van hat mennyiség meghatározására.)

Figyelembe véve a ϱ inkompatibilitási tenzor geometriai tartalmát a következőket mondhatjuk.

Ismert, hogy a beágyazó tér görbületi tenzorának eltűnését a felület pontjaiban, a felületre felépített koordinátarendszerben a felületelméletből ismert összefüggések fejezik ki. Nevezetesen, a *Gauss-féle Theorema Egregium* és a *Mainardi-Codazzi-egyenletek* (lásd LÁNCZOS [4] 184–188. old.). Ugyanakkor megmutatható, hogy a további három komponens eltűnése (a felületen, a felületre felépített koordinátarendszerben) a felület három alapformája közötti algebrai összefüggéssel ekvivalens. (Magát az algebrai összefüggést lásd pl. SULIKOVSKIJ [9] 252. old.) Meg kell jegyezni, hogy a *Theorema Egregium* azt is kifejezi, hogy a felületi (tehát a kétdimenziós) kovariáns deriváltak nem cserélhetők fel, a különbség a *Gauss-féle szorzatgörbület* (lásd pl. LÁNCZOS [4] 186. oldal).

Tekintsük a $\varrho n = 0$ egyenletet. Ez az egyenlet tenzoregyenlet, tehát tetszőleges koordinátarendszerben fennáll. Így a felületre felépített koordinátarendszerben is, ahol az n struktúrája $[0, 0, 1]$. Ennek megfelelően a $\varrho n = 0$ egyenlet mindössze ϱ három komponensének eltűnését írja elő. Ugyanakkor a felületre felépített koordinátarendszer alkalmazása miatt ez a három feltétel éppen a felületre vonatkozó *Theorema Egregium* és a két *Mainardi-Codazzi-egyenlettel* egyenértékű. Ezek fennállása — feltételezve az első alapforma pozitív definit voltát — a felületelmélet alaptétele értelmében (lásd pl. SULIKOVSKIJ [9] 262. old.) biztosítja, hogy a vizsgált felület első két alapformája (a merevtestszerű elmozdulástól eltekintve) egyértelműen határozza meg a felületet. Ezzel együtt a felületen a beágyazó tér görbületi tenzorának másik három (független) komponense, mint fentebb már utaltunk rá, a felület három alapformája közötti algebrai összefüggést adja vissza, azaz zérus értékű. Így a görbületi tenzor mind a hat független komponense eltűnik a felületen, tehát fennáll a $\varrho n = 0 \Rightarrow \varrho = 0$ implikáció is, ha

1°. a felületnek a megadott alakváltozási tenzormezőből számított első alapformája pozitív definit, valamint

2°. adott a felületnek a második alapformája.

A (3.4) *Kozák-féle peremfeltétel* önmagában nem tartalmazza ezt a két feltételt. Ezért (3.4)-ből nem következik (3.3), tehát (3.4) nem lehet szükséges és elégséges.

Vizsgáljuk meg a Kozák-féle peremfeltétel levezetését. Ez a *Stokes-tétel* alkalmazásán alapul. Az egyértékű eltolódásmező előállításának feltétele a *Cesaro-formulában* szereplő $\int_{M_0}^M B dr$ integrálnak az egyértékűsége. Ez egy felületen úgy is

megfogalmazható, hogy ott minden zárt h görbére fennáll a

$$(3.5) \quad 0 = \oint_h B \cdot t \, dr$$

összefüggés. Megjegyezzük, hogy a *Cesaro-formulában* szereplő B tenzor másodrendű (lásd pl. LUR'E [5] 63–64. old.). KOZÁK erre az összefüggésre (pontosabban transzponáltjára) alkalmazza a *Stokes-tételt*, és következményének a h görbe által bezárt A tartományon felírt

$$(3.6) \quad \iint_A (n \times \nabla) \cdot B^T \, dA = - \iint_A (B \times \nabla) \cdot n \, dA = 0$$

összefüggés-sort tekinti (KOZÁK [3] 230. old., számozatlan képlet), ahonnan a h görbék tetszőleges megválasztása és a

$$(3.7) \quad B \times \nabla = \varrho$$

azonosítás miatt már a fentemlített (3.4) összefüggésre jut.

A levezetés analíziséhez tekintsük a *Stokes-tételt* valamilyen X vektormezőre nézve, amely vektormezőre fennáll a

$$(3.8) \quad 0 = \oint_h X \cdot t \, dr$$

összefüggés. Alkalmazzuk a *Stokes-tételt*:

$$(3.9) \quad 0 = \oint_h t \cdot X \, dr = \oint_h dr \cdot X = \iint_A dA \cdot [\nabla \times X] = \iint_A n \cdot [\nabla \times X] \, dA.$$

Formailag — az $X = B$ helyettesítéssel — visszakaptuk KOZÁK eredményét: A negatív előjel nem játszik szerepet, valamint (3.9)-ben a transzponálást gyakorlatilag a skalárszorzat alkalmazása tartalmazza. De csak formailag kaptuk meg KOZÁK eredményét: (3.5) vektor-, (3.8) skalárétékű és — a *Stokes-tétel* csak vektormezőre értelmezett és nem tenzormezőre (lásd pl. KORN-KORN [2] 164. old.). Az, hogy a *Stokes-tétel* nem általánosítható tetszőleges módon, az a szakirodalomból ismert (pl. SZŐKEFALVI [10] 251. old., vagy STERNBERG [8], 3.3.5. tétel):

STOKES TÉTELE. Legyen M irányított m -dimenziós differenciálható sokaság. $D \subset M$ olyan reguláris tartomány M -ben, amelynek a \bar{D} lezártja kompakt, ω pedig $(m-1)$ -forma M -en. Ekkor fennáll a

$$(3.10) \quad \int_D d\omega = \int_{\partial D} \iota^*(\omega)$$

összefüggés, ahol ι jelöli a ∂D irányított részsokaság $\iota : \partial D \rightarrow M$ beágyazását M -be.

Itt arra kell rámutatni, hogy ha M , tehát D m -dimenziós sokaság, akkor ∂D $(m-1)$ -dimenziós, valamint ha ω $(m-1)$ -forma, akkor $d\omega$ m -forma. Ez összhangban van azzal, hogy a sokaságokon az integrálást úgy vezetjük be, hogy az m -dimenziós differenciálható sokaságon az m -ed-fokú ω differenciálformát integráljuk (lásd pl. [8] 3.2.6. definíciót, vagy [10] 247–248. old.).

A (klasszikus) *Stokes-tétel* kétdimenziós D tartományon értelmezett másodrendű differenciálforma (vektor rotációja) integrálját kapcsolja össze a tartomány egydimenziós ∂D határán értelmezett elsőrendű differenciálforma (a vektor komponenseivel meghatározott *Pfaff-forma*) integráljával.

Ennek megfelelően egyáltalán nem triviális, hogy a B másodrendű tenzorra vonatkozó (3.5) integrál átalakítható a *Stokes-tétel* segítségével. Ugyanis B nem is differenciálforma: nem alternál, de nem is alternálhat, mivel az egyedüli másodrendű differenciálforma az egydimenziós ∂D tartományon a 0 forma. Amennyiben a (klasszikus) *Stokes-tétel* általánosítható tetszőleges tenzorra, úgy annak bizonyítását célszerű lett volna mellékelni. Ugyanakkor az integrálmélet sokaságon, de a *Kozák-féle összeférhetőségi peremfeltétel* elégtelenségére vonatkozó megjegyzések is kétségesek teszik, hogy a *Stokes-tételnek* a fentebb feltételezett általánosítása megvalósítható lenne.

4. A szükséges és elégséges összeférhetőségi peremfeltételek

4.1. Mielőtt megadnók a szükséges és elégséges összeférhetőségi peremfeltételeket, egy lemmát ismertetünk.

LEMMA. Legyen az f folytonos függvény a V (egyszeresen összefüggő) tartományon azonos zérus értékű. Ekkor az f függvény a V tartomány S peremén is zérus értéket vesz fel.

A Lemma igazolásához az f függvény folytonosságát használjuk fel. Az f függvény folytonossága a $\bar{V} = V \cup S$ zárt tartományon azt jelenti, hogy $\forall x \in \bar{V} \forall (x_n) \subset \bar{V}$ olyan, hogy ha $x_n \rightarrow x$, akkor $f(x_n) \rightarrow f(x)$. Mivel $\forall x_n \in V$ -re $f(x_n) = 0$, az $x \in S$ választással nyerjük a Lemma állítását.

A Lemmát alkalmazhatjuk az inkompatibilitási tenzor folytonos komponenseire. Azaz, ha az inkompatibilitási tenzor valamely, illetve összes komponense eltűnik a vizsgált tartományon, akkor az, illetve mind eltűnik annak peremén is.

4.2. Ezután tétel formájában adjuk meg a szükséges és elégséges összeférhetőségi peremfeltételeket.

TÉTEL (a szükséges és elégséges összeférhetőségi peremfeltételekről). Tegyük fel, hogy az inkompatibilitási tenzor három alkalmasan kiválasztott komponense a vizsgált tartományon, míg a másik három komponense a tartomány peremén tűnik el:

$$(4.1) \quad \varrho^{XY} = 0, \quad \text{a } V\text{-n}$$

$$(4.2) \quad \varrho^{AB} = 0, \quad \text{az } S\text{-en,}$$

ahol az XY és AB három-három indexpár kiadja a teljes ij ($i = j = 1, 2, 3$) kombinációt (lásd KOZÁK [3], vagy BLOH [1] 314. old.). Akkor

- 1°. teljesül mind a hat összeférhetőségi mezőegyenlet,
- 2°. teljesül mind a hat összeférhetőségi peremfeltétel,
- 3°. teljessé válik a duál feladat peremfeltétel-rendszere.

A Tétel minden olyan $XY - AB$ particionálásra fordítható meg, amelyre az XY és AB indexpárok szerepet cserélhetnek.

A Tétel bizonyítását a 2. résszel kezdjük. A Lemma értelmében (4.1)-ből következik, hogy

$$(4.3) \quad \varrho^{XY} = 0, \quad \text{az } S\text{-en.}$$

Ezzel a 2. részt bebizonyítottuk.

A Tétel 1°. részének bizonyításához KOZÁK [3] után tekintsük az (3.2) *Bianchi-azonosság* bal oldalának és egy tetszőleges v vektornak a skalárszorzatát. Ezt integrálva a V tartományon, majd alkalmazva a *Stokes- (Gauss-Ostrogradszkij-) tételt*, az alábbi azonosságsort írhatjuk fel ($\partial V = S$):

$$(4.4) \quad 0 = \int_V v_j \nabla_i \varrho^{ij} dV = \int_S v_j \varrho^{ij} n_i dS - \int_V (\nabla_i v_j) \varrho^{ij} dV.$$

A felületi integrál a 2. rész értelmében zérus.

A térfogati integrálon belüli összeget XY és AB indexpárok szerint csoportosítjuk:

$$(4.5) \quad \int_V \sum_{(XY)} (\nabla_X v_Y) \varrho^{XY} dV + \int_V \sum_{(AB)} (\nabla_A v_B) \varrho^{AB} dV = 0.$$

Az első integrál (4.1) értelmében zérus, a második integrál kiszámításához felhasználjuk ϱ szimmetriáját, így mindössze három tagból álló kifejezést kell integrálni. Továbbá ϱ^{AB} „együtthatói” a ∇u gradienstenzor szimmetrikus része, illetve annak is csak az AB indexpárok szerinti komponensei. Az $XY - AB$ indexpár „megfelelő” kiválasztása azt jelenti, hogy, v tetszőlegessége folytán, ϱ^{AB} „együtthatói” tetszőleges függvényeknek választhatók (KOZÁK [3]). WASHIZU [11] nyomán az

$$(4.6) \quad \frac{1}{2}(\nabla_A v_B + \nabla_B v_A) = \varrho^{AB}$$

választással élünk, így (4.5) az

$$(4.7) \quad \int_V \sum_{(AB)} (\varrho^{AB})^2 dV = 0$$

alakra hozható, ahonnan csak a

$$(4.8) \quad \varrho^{AB} = 0, \quad \text{a } V\text{-n}$$

megoldások lehetnek. Ezzel az 1°. részt bebizonyítottuk.

A Tétel 3°. részének bizonyításához elegendő összeszámlálni a peremfeltételeket, amennyiben pontonként három peremfeltételt ad az előírt eltolódás és/vagy felületi megoszló erő és hármat ad a (4.2) előírt összeférhetőségi peremfeltétel.

A szükségesség és elégségesség abból következik, hogy a (4.1,2) feltételrendszerből igazoltuk a (4.3) és (4.8) fennállását, amelyből az azonos struktúra miatt igazolható az (4.1,2) teljesülése. Ez alól az az eset a kivétel, amikor az AB indexpár-hármas nem tekinthető alkalmasan kiválasztottnak. Erre példa a *Descartes-féle koordinátarendszerben* az $AB := 11, 12, 13$ indexpár-hármas, és az 1, 2 és 3 ciklikus felcserélésével nyert további két eset. Ezzel igazoltuk a Tétel megfordíthatóságát, valamint a Tétel megfordíthatóságára vonatkozó feltételt is.

Ezzel a Tételt bebizonyítottuk.

A Tételből következik, hogy WASHIZU — az általa tárgyalt két speciális esetben ($XY = (11, 22, 33)$, $AB := (12, 23, 31)$ és duálisa) adott — megoldása nemcsak szükséges, hanem elégséges is. Együttal bebizonyítottuk azt is, hogy sejtése igaznak bizonyult, a (2.9) feltétel valóban redundáns.

4.3. Vizsgáljuk meg a *Kozák-féle összeférhetőségi peremfeltételt*. Ugyanis a fent bizonyított Tételből következik, hogy a *Kozák-féle összeférhetőségi peremfeltétel* nem szükséges. Ezzel szemben KOZÁK a (2.11) és (2.10) kiinduló feltevések mellett bizonyítja (2.12)-t (lásd [3]-at), azaz (4.8)-at, és ennek során a Lemmát (azaz ϱ folytonosságát) nem használja ki. Ezért a bizonyításhoz — a (4.4) integrálatalakításból látható, hogy — szüksége van a (2.10) alatti összeférhetőségi peremfeltételre. Ugyanakkor a Lemma fennáll, ezért a (2.10) *Kozák-féle összeférhetőségi peremfeltételeket* — a (4.4) integrálatalakításban szereplő felületi integrál eltűnéséhez — elegendő a

$$(2.10') \quad \sum_{(AB)} \varrho^{AB} n_B = 0, \quad \text{az } S\text{-en}$$

alakban felírni. (Ennek igazolásához elegendő (4.4)-be behelyettesíteni a (4.1) és a Lemma következményeként fennálló (4.3) összefüggést, illetve a (2.10') módosított *Kozák-féle összeférhetőségi peremfeltételt*.)

Ugyanakkor a *Kozák-féle összeférhetőségi peremfeltétel* módosított, (2.10') alakja lehetővé teszi a kétféle — tehát a (4.2) és a (2.10) — összeférhetőségi peremfeltétel ekvivalenciájának kimutatását. Az, hogy (4.2)-ből következik (2.10'),

és így a (4.1) valamint a Lemma miatt (2.10), az triviális. A fordított állítás igazolásához tekintsük a (2.10') alatti összefüggést mint a három ϱ^{AB} mennyiségre nézve egy homogén, lineáris egyenletrendszer. Ha az n_B mennyiségekből alkotott együtthatómátrix determinánsa nem zérus, akkor a (2.10') egyenletrendszernek csak a $\varrho^{AB} = 0$ triviális megoldásai léteznek (lásd pl. KORN-KORN [2] 38. old.), ami ekvivalens (4.2)-vel. Ha viszont a fent bevezetett determináns értéke zérus, akkor a független egyenletek száma nem lehet három (csak annál kevesebb, lásd a fenti ir. hiv.-t). Ekkor ugyan a (2.10) összefüggés a (4.4) integrálalakítás során — a (4.3)-mal együtt — „lenullázza” a felületi integrált, de nem ad elegendő számú előírt összeférhetőségi peremfeltételt.

(Meg kell vizsgálni, hogy vajon ez az utóbbi eset fennállhat-e. Igen. Két okból is. Az egyik, ha az adott koordináta-rendszerben az n egységvektor valamelyik bázisvektorral párhuzamos. A másik, ha *görbevonali* koordináta-rendszerben az $XY := (11, 12, 13)$, illetve az 1, 2 és 3 ciklikus felcserélésével nyert indexpár-hármas választással élünk.)

Összefoglalva, ha a (2.10') módosított Kozák-féle összeférhetőségi peremfeltétel egyértelműen megoldható a ϱ^{AB} mennyiségekre nézve, akkor (2.10') — azaz (2.10) — ekvivalens (4.2)-vel, és így (2.10')-t *nem szükséges* alkalmazni, ha nem oldható meg egyértelműen a ϱ^{AB} mennyiségekre nézve, akkor nem szolgáltat elegendő összeférhetőségi peremfeltételt, és így (2.10')-t *nem lehet* használni, csak (4.2)-t.

Tehát a Lemmát, azaz a ϱ folytonosságát felhasználva egyúttal azt is megmutattuk, hogy (a fentebb, a zárójelbe tett esetektől eltekintve) KOZÁK általános formában megfogalmazott megoldása a speciális esetekben visszaadja WASHIZU megoldásait, ha a Kozák-féle peremfeltételt a (2.10') alatti módosított alakban írjuk fel.

4.4. A fenti fejtegetések ismételten rámutatnak arra, hogy meg kell különböztetni a felületen előállítható egyértékű elmozdulásmező szükséges és elégséges összeférhetőségi feltételét a kontinuummechanika duál peremérték-feladata szükséges és elégséges összeférhetőségi peremfeltételétől.

Az első $\varrho = 0$ (a felületen) alakban írható fel.

A második a $\varrho^{AB} = 0$, vagy a $\sum_{(AB)} \varrho^{AB} n_B = 0$ (a felületen), egymással ekvi-

valens alakban írható fel. (Az n_B -re vonatkozó megszorítások megismétlésétől itt eltekintünk.)

Végül rá kell mutatni arra is, hogy a kontinuummechanika duál peremérték-feladata szükséges és elégséges összeférhetőségi peremfeltétele nem önmagában szolgáltatja a peremen (azaz egy felületen) előállítható egyértékű elmozdulásmező szükséges és elégséges összeférhetőségi feltételét, hanem az előírt összeférhetőségi mező-egyenletekkel és az inkompatibilitási tenzor folytonosságával együtt. (Ez a Tétel 2°. részének állítása.) Ezért nem szükséges, hogy a kétféle szükséges és elégséges összeférhetőségi feltétel egybeessen.

4.5. Végül teszünk egy megjegyzést a Tétel 1°. és 2°. részének függetlenségére a Descartes-féle koordináta-rendszerben.

A *Descartes-féle koordináta-rendszerben a Bianchi-azonosságot* (4.1) miatt a

$$(4.9) \quad \sum_{(AB)} \nabla_A \varrho^{AB} = 0, \quad \text{a } V\text{-n}$$

alakban is felírhatjuk, mivel $\nabla_A = \partial/\partial A$ ($A = x, y, z$). (Lásd még WASHIZU [11] bizonyítását a *Morera-féle feszültségfüggvény* esetében.) Ekkor a (4.4) helyett elegendő a

$$(4.10) \quad 0 = \int_V \sum_{(AB)} v_B \nabla_A \varrho^{AB} dV = \int_S \sum_{(AB)} v_B \varrho^{AB} n_A dS - \int_V \sum_{(AB)} (\nabla_A v_B) \varrho^{AB} dV.$$

azonosságot felírni. A felületi integrál (4.2) értelmében zérus, innét a fenti bizonyítást alkalmazva igazoljuk (4.8) teljesülését. A lényegi eltérés az, hogy (4.8) igazolásához csak (4.1) és (4.2) feltételeket használtuk fel és (4.3)-mat nem. Azaz a *Descartes-féle koordináta-rendszerben* a Tétel 1°. és 2°. része egymástól függetlenül is bizonyítható. Ezért tudta WASHIZU (*Descartes-féle koordináta-rendszerben*) a Lemma nélkül is a *Southwell-paradoxont* feloldani.

4.6. A *Southwell-paradoxon* és feloldása kapcsán a következőket kell megemlíteni.

A *Southwell-paradoxon* tulajdonképpen a lineáris rugalmasságtan duál feladata variációs elveinek felírása során „lépett fel” (SOUTHWELL [6,7]), az eljárás korrektségét elsőnek WASHIZU [11] igazolta. A paradoxon tetszőleges görbevonali koordináta-rendszerbeli feloldása (KOZÁK [3]) teszi lehetővé a duál feladat variációs elveinek teljessé tételét, illetve a primál feladat variációs elveivel való teljes analógiának a felállítását. Jelen dolgozatban pedig egyrészt megadtuk a szükséges és elégséges összeférhetőségi peremfeltételeket, másrészt rámutattunk arra, hogy a módszer alkalmazható tetszőleges nemlineáris alakváltozás-eltolódás összefüggés esetén is.

4. Összefoglalás

A tanulmányban a *Southwell-paradoxon* megoldása során felmerült összeférhetőségi peremfeltételekkel foglalkozunk.

Bebizonyítottuk, hogy ha az inkompatibilitási tenzor három megfelelően kiválasztott komponense kielégíti az összeférhetőségi mezőegyenleteket, a másik három az összeférhetőségi peremfeltételeket, akkor

1°. ez szükséges és elégséges feltétele annak, hogy a másik három összeférhetőségi mezőegyenlet is teljesüljön, továbbá, hogy

2°. ez szükséges és elégséges feltétele annak, hogy a másik három összeférhetőségi peremfeltétel is teljesüljön.

3°. Végül a feltételként kirótt három összeférhetőségi peremfeltételt a primál rendszerben megszokott peremfeltételekhez csatolva a duál rendszer szükséges hat peremfeltételéhez jutunk.

A tanulmányban ezzel együtt bebizonyítottuk azt is, hogy a WASHIZU által alkalmazott összeférhetőségi feltétel nemcsak szükséges, hanem elégséges is.

Rámutattunk arra is, hogy fenti megoldás alkalmazható tetszőleges nemlineáris alakváltozás-eltolódás összefüggés esetén is.

IRODALOM

- [1] БЛОХ (БЛОХ, В. И.), *Теория упругости* (Изд. Харьковского Унив., Харьков, 1964).
- [2] KORN, G. A. és KORN, T.M., *Matematikai kézikönyv műszakiaknak* (Műszaki Könyvkiadó, Budapest, 1975).
- [3] KOZÁK, I., „A lineáris elasztostatika feszültségekkel felírt mezőegyenleteiről és peremfeltételeiről”, *Műszaki Tudomány*, vol. 57 (1979), 423–446.
- [4] LÁNCZOS, K., *A geometriai térfogalom fejlődése* (Gondolat, Budapest, 1976).
- [5] ЛУР'Е (ЛУРЬЕ, А. И.), *Теория упругости* (Наука, Москва, 1970).
- [6] SOUTHWELL, R. V., „Castigliano's Principle of Minimum Strain Energy”, *Proc. Roy. Soc. A* **154** (1937), 4–21.
- [7] SOUTHWELL, R. V., „Castigliano's Principle of Minimum Strain Energy and the Condition of Compatibility for Strain”, *S. Timoshenko, 60th Anniversary Vol. 1938. The MacMillan Co.*, 211–217.
- [8] STERNBERG, S., *Lectures on Differential Geometry* (Prentice Hall, N. J., 1964).
- [9] СУЛИКОВСКИЙ (ШУЛИКОВСКИЙ, В. И.), *Классическая дифференциальная геометрия* (ФМ, Москва, 1963).
- [10] SZŐKEFALVI (SZŐKEFALVI-NAGY GY., GEHÉR L. és NAGY P.), *Differenciálgeometria* (Műszaki Könyvkiadó, Budapest, 1979).
- [11] WASHIZU, K., „A Note on the Conditions of Compability”, *J. of Math. and Physics* **36** 4 (1958), 306–312.

(Beérkezett: 1990. március 2.)

(Átdolgozva beérkezett: 1991. november 12.)

LÁMER GÉZA
1126 BUDAPEST, XII. KISS J. ALTGY. U. 46.

DETERMINATION OF NECESSARY AND SUFFICIENT COMPATIBILITY BOUNDARY CONDITIONS

G. LÁMER

In the paper the compatibility boundary conditions arising in the solution of *Southwell-paradoxon* are investigated. There are given the necessary and sufficient compatibility boundary conditions. We prove that the solution given by WASHIZU is not only necessary but also sufficient.

A RÁKKUTATÁS MATEMATIKAI ALAPJAI*

TUSNÁDY GÁBOR

Budapest

Előszó

1982 nyarán az MTA Matematikai Kutatóintézete levelet írt az Országos Onkológiai Intézet igazgatójának, Eckhardt Sándornak, hogy a két intézet együttműködését kezdeményezze. Válaszul novemberben a rákkutatók egy népes csoportja ellátogatott az intézetünkbe, és felvázolták azokat a területeket, ahol megítélésük szerint matematikai segítségre szorulnak. E látogatás következtében rendszeres közös munka alakult ki a két intézet között, amelynek felgyorsítására 1989-ben két évre OTKA-támogatást kaptunk. Ez a beszámoló elsősorban az 1989–90 évek eredményeit ismerteti, de utalunk terveinkre is és a korábbi tapasztalatokra is.

Időközben az Országos Onkológiai Intézetnek saját számítástechnikai bázisa alakult ki, amelynek üzemeltetéséhez sok matematikai kérdést kellett megoldani. Ezekről itt nem fogunk beszélni, a két intézet együttműködése elsősorban a kutatás fejlesztését szolgálja, így ez a két terület az évek során természetes módon elhatárolódott, noha közöttük nyilván vannak érintkezések és átfedések.

Munkánk során az Országos Onkológiai Intézet dolgozói részéről fáradhatatlan segítséget kaptunk. A legszorosabb együttműködés a következő kutatókkal alakult ki: Szentirmay Zoltán, Holzinger László, Bakács Tibor, Ringwald Gábor, Baki Márta és Hullán Lehel.

Együttműködésünk egyik célja az volt, hogy a téma iránt érdeklődő matematikusok munkáját irányítsuk. Ennek megfelelően a munkába hosszabb-rövidebb időre sokan bekapcsolódtak, közülük a legfontosabbak a következők voltak: Vincze István, Pick Róbert, Fritz József, Rejtő Lídia, Varga László, Arató Miklós, Michaletzky György, Tóth János, Bártfai Pál, Major Péter, Csiszár Imre, Komlós János, Paul Shields, Bornemisza Eszter és Magyarfi István.

A továbbiakban az Országos Onkológiai Intézetet röviden csak „klinika”-ként említjük, utalva annak beteggondozó jellegére.

*A dolgozat a 1818 számú OTKA kutatás beszámoló jelentése.

Bevezetés

Matematikát a rákkutatásban a következő területeken lehet alkalmazni:

- a rákos folyamatok megértése és leírása;
- az orvos segítése a diagnózis kialakításában;
- a gyógyítás és kezelés támogatása;
- a betegek sorsának követése és ezen keresztül az előző pontok eredményeinek ellenőrzése és fejlesztése.

Egy gyakorló kutató a munkája során természetesen a legváltozatosabb matematikai és számítástechnikai eszközöket veszi igénybe. Rendszerezhetnénk az érintkezési területeket matematikai szempontból is, a két felosztás közti választás szubjektív. Mi elsősorban az alkalmazhatóság és a későbbi fejlemények előkészítése érdekében követjük a fenti sorrendet. Látni fogjuk, hogy sok helyen az eredmények hézagosak, még csak homályos, kialakulatlan elképzeléseink vannak. Talán nem lesz érdektelen itt azokat is elmondani, hiszen hazánkban a mi munkánk az első ilyen jellegű átfogó tevékenység.

A rákos folyamatok megértése és leírása

A rákos folyamatok biológiai leírását legrészletesebben James D. Watson A gén molekuláris biológiája című könyvében találjuk meg, amelynek újabb angol kiadása már részletesebb mint az, amelyiket magyarra is lefordítottak, és a könyv körül kialakult szerzőgárda napjainkban is folyamatosan dolgozik a további kiadások előkészítésén.

Matematikailag először is a sejt-automaták elméletét kell említeni a rákkal kapcsolatban. Egy sejt-automatának inputja, állapota és outputja van, ezek általában egy-egy véges halmaz elemei lehetnek. Működhet az automata folytonos vagy diszkrét időben. Magát az automata működését az a szabály határozza meg, ahogyan az input hatására az állapot megváltozik, és esetleg output keletkezik. Ezekből az automatákból hálózatokat lehet építeni, amelyben az outputok valamilyen struktúra szerint a hálózat meghatározott elemei számára inputtá alakulnak.

Az eredetileg merev, determinisztikusan működő hálózat különböző helyein az idők során megváltozhat. Eleve az is beépíthető a modellbe, hogy a hálózatnak új szemei, automatái épüljenek ki, amelyek keletkezésük függvényében szervesen csatlakoznak a meglévő rendszerhez, és más szemek közben elhalnak. Ez az önreprodukció függhet az állapottól vagy a véletlentől. Stabilisnak mondunk egy rendszert, ha enyhe véletlen hatásokkal szemben ellenálló, eredeti térbeli homogenitását időben képes megőrizni.

Előfordulhat, hogy egy rendszernek több stabilis állapota van. Lokálisan átbil-lenhet egyikből a másikba, és a különböző stabilis állapotok szigetekké csoportosulhatnak, amelyek kölcsönösen meg akarják egymást szüntetni.

Az élő szervezet sejtjeinek az állapotát a DNS határozza meg. Emberre a lehetséges állapotok száma igen nagy. Egy egész ember feltérképezése még valószínűleg több évtized munkája lesz. Akár egyetlen sejté is. Elvileg egy ember sejtjeinek DNS-tartalma identikus kellene legyen, hiszen egyetlen sejt másolatai. Arról az egyetlen sejtről diadikus fa ágai szerint keletkeznek a másolatok, az ágak szerinti összetartozó lomb-darabok funkcionálisan összetartoznak. A különböző ágak nem egyforma hosszúak, az élet során a különböző szövetek reprodukciója más és más ütemű. Általában a hiány diktálja az újratermelést, és az egyes sejtek a membránon, a sejtek falán át szereznek tudomást a külvilágról. Így állnak egymással is kapcsolatban, elsősorban a testen belüli különböző nedvek keringése révén.

Az időben lineárisan egymást követő osztódások során másolási hibák keletkezhetnek. E hibák többsége kijavítódik. Vannak eleve életképtelen másolatok, mutációk, másokat a környezet blokkol. De e folyamatok közben alapvető kérdés az, hogy ki az úr a házban, ki képviseli az igaz utat, az egészséget.

Általános tapasztalat szerint a rák alattomosan, észrevétlenül kezdődik. Olyan mutációk állhatnak mögötte, amelyek csak hosszú távon, sok-sok egymást követő osztódás után fejtenek ki észrevehető funkcionális változást, ha már elég sok új típusú sejt keletkezett. Többségükben ezek az új sejtek kevésbé differenciáltak, arra hasonlítanak, ami a csecsemők szervezetében található. Lassan kezdjük megtanulni a DNS-ben azokat a helyeket, ahol a leggyakoribbak a rákra vezető mutációk. A fejlődési, osztódási lánc mentén nem csak mutációk keletkeznek. Előfordulhat, hogy csak a sejtmag osztódik, a sejt nem. Ezt követően a kétszeres DNS-állományú sejt osztódik tovább, újra szabály szerint, és az ilyen kétszeres DNS-tartalmú sejtekből külön kolóniák keletkezhetnek. Ez a lépés újra megismétlődhet, négyszeres, nyolcszoros DNS-tartalmú sejtek is keletkezhetnek. Nem mindig szabályos a duplázódás, van amikor törmelékek keletkeznek, és azok sokszorozódnak tovább szabályos osztódások révén.

A DNS-lánc eredeti sorozatában tehát egy-egy elem kicserélődhet, törölődhet, vagy beszűrődhet. Hosszabb darabok is kieshetnek, megismétlődhetnek, vagy átrendeződhetnek. Részben a DNS-láncok belső struktúrájának vizsgálata és különböző sorozatok hasonlóságának a mérése, részben a DNS-ből dekódolt fehérjék hasonló vizsgálata vezetett egy új tevékenység kialakítására, amit angolul ALIGNING-nek neveznek, magyarul talán felsorakoztatásnak mondhatunk.

Égészen általánosan azt mondhatjuk, hogy ha valamely ABC feletti sorozatokon azokat egy másik ABC feletti sorozatba vivő értékelt transzformációk serege áll rendelkezésre, akkor a két ABC feletti sorozatok között távolságot létesíthetünk, ha azt a minimális költségű transzformáció költségeként definiáljuk. (A sorozatok eltérését „távolságnak” nevezzük ugyan, ezzel azonban nem állítjuk, hogy ez a mennyiség bármilyen matematikai távolság tulajdonságaival rendelkezze.)

Ha a második ABC felett több sorozat adott, azok közös forrásból való származását ellenőrizhetjük úgy, hogy megkeressük azt az első ABC-ből felírt sorozatot, amelytől mérve az adott sorozatok távolságait, azok összege minimális.

E távolságok meghatározására léteznek algoritmusok, de azok egzakt végrehaj-

tása ma még csak kis méretben lehetséges: vagy csak két sorozatot szabad vizsgálni, vagy több, de nagyon rövid sorozatot. Heurisztikus algoritmusokat mi is dolgoztunk ki, ezek futási tapasztalatai jónak mondhatóak. De mindaz, amit ebben a fejezetben elmondtunk, még csak az előkészítés állapotában van. Valódi felhasználásának sok technikai akadálya van még, de célszerűnek látjuk a matematikai háttérrel időben előkészíteni, nehogy a technika megfelelő fejlődése után a szükséges matematika megszerzése vegyen el időt a betegek gyógyításától.

Az orvos segítése a diagnózis kialakításában

Ma a diagnózis legfontosabb támasza a szövettani vizsgálat. A szervezetből kivett metszetek leggyakoribb vizsgálata optikai úton történik. Mi a legtöbbet a következő feladattal foglalkoztunk. Egy berendezéssel párhuzamos egyenesek mentén meghatározható az egyes egyenesek sejtmagon belüli darabjának a hossza. Ha a választott egyenesek a rájuk merőleges síkot egy négyzetrács pontjaiban metszik, az egyes sejtmagokra vonatkozó mérések egy mátrixba foglalhatóak össze. Tipikusan ez a mátrix 20×30 méretű, de alakja és mérete függ a sejtmag nagyságától és alakjától.

A gyakorlatban a mátrix elemeit összeadják, ezt tekinthetjük a sejtmag térfogatának a becslésének. De a következő megfontolások szempontjából mindegy is, hogyan mérjük a sejtmag tömegét. Az osztódási ciklust a következő részekre szokás bontani:

- G_0 fázis, a sejt nem osztódik,
- G_1 fázis, a sejt osztódni kezd, de a sejtmag kettőzése még nem kezdődött el,
- S fázis, a sejtmag anyaga megkétszereződik,
- G_2 fázis, a sejtmag tömege kétszeres, egyéb komponensek kétszereződnek,
- M fázis, a sejt két sejtté osztódik.

Mi a továbbiakban nem fogjuk a G_0 és G_1 valamint a G_2 és M fázisokat megkülönböztetni. Ezekben ugyanis a sejtmag tömege ugyanaz, és minket most a sejtmag tömegének az eloszlása érdekel.

Jelöljük a G_0 és G_1 fázisban a sejtmag tömegét valamilyen virtuális egységben A -val. Homogén sejtpopulációban ez minden egyes sejtre elvileg ugyanaz az érték. Hasonlóan konstans a G_2 és M fázisban a sejtmag tömege, jelöljük ezt B -vel, természetesen általában $B = 2A$. Jelöljük végül az S fázisban a sejtmag tömegét C -vel. Szemben A -val és B -vel C nem állandó, hanem időben változó érték A és B között.

Ha egy adott pillanatban egy szövet sejtjeit vizsgáljuk, egy véletlenszerűen kiválasztott és S fázisban levő sejt magjának a tömege a véletlentől függ, valószínűségi változó, amelynek meghatározott eloszlása van. Ez az eloszlás három tényezőtől függ:

- a sejtmag termelésének a sebessége, ezt állandónak vehetjük, így e tényező önmagában ún. egyenletes eloszlásra vezet;

- a sejtpopuláció szinkronitása, vagyis az egyes sejtek cikluson belüli helyzetének megegyezése vagy aszinkronitása vagyis e helyzetek függetlensége;
- a sejtpopuláció növekedési üteme, amely egyrészt a különböző fázisok idejének az arányától, másrészt a sejtek kipusztulási valószínűségétől függ.

Vizsgáljuk meg ezt a harmadik tényezőt. Az egyszerűség kedvéért tegyük fel, hogy a fázisok ideje állandó, jelöljük ezeket rendre T_0 -lal, T_1 -gyel, T_S -sel, T_2 -vel és T_M -mel, és tegyük fel, hogy kettéosztódás után mindkét leánysejt egymástól függetlenül k valószínűséggel folytatja az életét, és $1 - k$ valószínűséggel elpusztul.

Jelöljük a t időpillanatban az u -nál kisebb tömegű sejtmaggal rendelkező sejtek számát $N(t, u)$ -val, $N(t, u)$ várható értékét $F(t, u)$ -val. Tegyük fel, hogy a populáció aszinkron, vagyis $F(t, u)$ deriválható u szerint:

$$F(t, u) = P(t) + \int_A^u f(t, v) dv, \text{ ha } A < u < B,$$

és legyen

$$Q(t) = \int_A^B f(t, v) dv,$$

$$F(t, u) = P(t) + Q(t) + R(t), \text{ ha } u > B,$$

vagyis $P(t)$, $Q(t)$ és $R(t)$ rendre legyen az A -val egyenlő, A és B közötti, és B -vel egyenlő tömegű sejtmaggal rendelkező sejtek számának a várható értéke. Akkor

$$P'(t) = -f(t, A) + 2kf(t - T_B, B),$$

$$R'(t) = f(t, B) - f(t - T_B, B),$$

ahol $T_A = T_0 + T_1$, és $T_B = T_2 + T_M$, és ha a magtermelés egyenletes sebességű, akkor $f(t, u)$ csak $(t - u)$ -tól függ. Ennek a függvényegyenletnek azt a megoldását keressük, amelyben az

$$f(t, u)/(P(t) + Q(t) + R(t))$$

hányados nem függ t -től, és ezt az állandó, csak u -tól függő függvényt nevezzük stacionárius megoldásnak.

Meggondolásunk lényege, hogy lehetővé teszi azt, hogy változó méretű, növekedésben levő sejtpopuláció esetében stacionárius eloszlásról beszéljünk. Belátható, hogy ez a stacionárius megoldás exponenciális függvény, ezért mi az S -fázisbeli eloszlást exponenciális függvénynek választjuk.

Jelöljük valamely sejtpopuláció véletlenszerűen választott tagjában a sejtmag tömegét virtuális egységben mérve X -szel. Ez a mennyiség csak valamilyen, általában additívnek feltételezhető Z zajjal mérhető, tehát a mérés eredménye az

$$Y = X + Z$$

valószínűségi változó. Egyes feltételezések szerint a mérési hiba arányos a mért mennyiséggel, ebben az esetben

$$Y = X(1 + Z).$$

Mi a Z változót normális eloszlásúnak tételezzük fel.

A feladat tehát az, hogy a hibára tett konkrét feltételek mellett az Y változóra vonatkozó

$$Y_1, Y_2, \dots, Y_N$$

független, egyforma eloszlású minta alapján meghatározzuk az X változó eloszlását, vagy annak paramétereit. A legegyszerűbb esetben a három fázis P , Q , R valószínűsége, az A mennyiség és az S -fázisra feltételezett exponenciális eloszlás paramétere ismeretlen. Ez utóbbit előre meghatározott állandónak feltételezve az ismeretlenek száma 3-ra redukálható, hiszen $P + Q + R = 1$. Negyedik ismeretlen a zaj szórása lehet, amely szintén választható külső konstansnak is.

A diagnózis szempontjából a legfontosabb P értéke, hiszen ha ez kicsi, a populáció gyorsan osztódik. Ennél is fontosabb azonban azt eldönteni, hogy egyetlen homogén populációból áll-e a vizsgált anyag, vagy sem. Általában ugyanis X eloszlása különböző A alapmennyiségekhez tartozó eloszlások keveréke. Ebben az esetben meghatározandó a komponensek száma, n , a hozzájuk tartozó A_1, A_2, \dots, A_n nyugalmi sejtömegek, és a komponensek $\pi_1, \pi_2, \dots, \pi_n$ aránya. A klinikai tapasztalat szerint a diagnózis szempontjából a legfontosabb n értéke, vagy valamely más, a populáció inhomogenitását mérő mennyiség.

Statisztikai szempontból a feladat meglehetősen nehezen kezelhető. Önmagában a becslések megtalálása bonyolult, lassan konvergáló eljárásokat igényel, és sajnos a becsléseket definiáló követelmények általában nem határozzák meg egyértelműen magukat a becsléseket. A becslések kialakítása sok mellékkörülménytől függ, amelyek együttes megválasztása, az eljárás „behangolása” alapos felkészültséget követel. Ennek ellenére sikerült összefüggő programrendszert kidolgozni a feladat megoldására, és a kérdés elméleti vizsgálata terén is vannak eredményeink. De a kutatás egyáltalán nincs lezárva, mondhatni a legfontosabb kérdések még megválaszolatlanok.

Visszatérve a bevezetőben tárgyalt mérésre, természetes gondolat a sejtmag kétdimenziós vetületét meghatározó teljes mátrixot felhasználni a diagnózis céljaira. Ehhez diszkriminatív statisztikákra van szükség, amelyeket csak hosszú időn át együttműködő csoport képes kialakítani. Ma már természetesen vannak az ilyen feladatokra tervezett automatikus eljárások is, amelyek közül eggyel a klinika is dolgozik évek óta. A kérdés elválaszthatatlan a műszaki feltételektől, amelyek egyrészt a képfelbontás finomságát, másrészt a diszkriminációt végző matematikai eljárás bonyolultságát határozzák meg.

Mi kezdetben magát a kétváltozós függvényt közelítettük, amelyik a felületet leírja. Csak lassan derült ki a számunkra, hogy ez az út lényegében járhatatlan: egyrészt nagyobb a szóbanforgó felületek változékonysága, semhogy reálisan paraméterezhető felület-családdal azt követni lehetne, másrészt igen valószínű, hogy nem a konkrét alak, hanem annak simasága diszkriminatív.

Napjaink egyik leghatásosabb diszkriminációs módszere az ún. leghatásosabb vetület módszer (projection pursuit), aminek a mi esetünkre való alkalmazásaként tipikus vetület-profilokat határoztunk meg, és minden egyes sejtmag-mátrixhoz ezek

közül kerestük meg a legjobban közelítőket. Ezzel a módszerrel érzékeny, több lépcsős sejt-populáció leírást kapunk, amely a típusgörbék gyakoriságából, összetételéből, és az azok által adott közelítés hibájából alakítja ki az adott populáció diagnózisát.

A rák arca nagyon sokféle, olyasmi a rák, mint a veleszületett rendellenességek. Közös az eredet, a szabályok, de attól függően, hogy melyik szervet támadja meg, milyen a fejlődés dinamikája, rengeteg változata van. Újabban lehetőség nyílt a DNS bizonyos pontjainak az azonosítására, kétségtelen, hogy ennek az eljárásnak a legnagyobb a fejlesztési lehetősége. Az élet során a szervezetben folyamatos a sejtosztódás. Ma még nem ismerjük az itt keletkező hibák dinamikáját. Sem azt, hogy a föld rákbetegeiben előforduló DNS-szekvenciák végül is mekkora szórást mutatnak, mint ahogy azt sem tudjuk, hogy az egészséges emberek szekvenciáinak mekkora a szórása. Hogy nagy, azt a szervátültetések sorsa mutatja. E finom módszerek ismeretében az ember bizonyos türelmetlenséggel gondol a durvábbakra. Viszont az orvosi gyakorlat számtalan területen azt mutatja, hogy a kutatás valamely igen részletes módszere nem feltétlenül szorítja ki a régi, egyszerű, jól bevált módszereket, például a röntgen nem szüntette meg azt, hogy az orvos egyszerűen a beteg hátához nyomja a fülét.

A gyógyítás és a kezelés támogatása

A szervezet védekező rendszere az immunrendszer, amelynek vizsgálatára a klinikán a következő laboratóriumi mérést alkalmazzák. Félreértések elkerülése végett megjegyezzük, hogy az ismertetendő folyamat a szervezetben nem megy végbe, csak laboratóriumi körülmények között jön létre. Különböző mennyiségű antitesttel szenzitivizált vörös véresejtet vagy fehér véresejttel, vagy annak a hatóanyagával, a komplementtel pusztítanak el. Az elpusztított vörös véresejtek száma függ az időtől, a vörös véresejtek összes számától, az antitest-koncentrációtól és a pusztítást végző anyag, közös néven effektor mennyiségétől. Első feladatunk e függvény meghatározása volt.

Ennek a feladatnak a megoldására három különböző módszert alkalmaztunk. Egyrészt a legkisebb négyzetek módszerével különböző többváltozós függvénnyel közelítettük a mérési eredményeket. A legegyszerűbb és leghatásosabb eljárásnak a változók logaritmikus transzformációja utáni lineáris regresszió bizonyult. A második módszer az irodalomban Breiman–Friedman ACE-módszer néven ismert eljárás módosítása volt, amit mi görbevonalú szórásanalízisnek neveztünk el. A módszer lényege, hogy magát a végeredményt és az egyes komponenseknek arra kifejtett hatását ismeretlen, nem paraméteres függvénynek tekinti, és ezekből az egyváltozós, a mérési eredmények által meghatározott függvényekből építi fel a keresett többváltozós függvényt. Mi az ismeretlen függvényeket monotonnak tételeztük fel, és ezáltal lehetővé tettük a monoton regresszióra kidolgozott effektív numerikus eljárások alkalmazását. A harmadik módszer a mérési eredmények közelítésére a

folyamat dinamikáját tükröző differenciálegyenletek keresése volt. Ezt a módszert elsősorban az ún. kompetíciós mérések kiértékelésére alkalmaztuk.

A laboratóriumi gyakorlatban legtöbbször A és B vörös vérsejtet (közös néven targeteket) használnak, és évek óta folyik annak a vizsgálata, hogy ha az effektor ezeket egyidejűleg pusztíthatja, akkor a kétféle targetre kifejtett hatás hogyan függ azok szenzitivizálásától. Ha ugyanis a kétféle target antitesttel való fedettségét egymástól függetlenül növeljük, akkor egyrészt természetes módon irányítani lehet az effektort, az ugyanis várhatóan azt a targetet pusztítja jobban, amelyiken több antitest van, másrészt a két folyamat között kölcsönhatás alakulhat ki, bizonyos antitest-arányok mellett a két folyamat serkentheti, máskor meg gátolhatja egymást. A statisztikus feladata azt eldönteni, így van-e ez, és ha igen, mi lehet az oka.

Magát a vizsgált folyamatot már elég jól ismerjük. Különösen ami a komplex reakciót illeti, itt kilenc komponens van, és ezek kialakulása, kölcsönhatása is ismert. Nem ismertek azonban a folyamatok paraméterei, és az, hogy a keletkezett termékek csak annak a sejtnek az elpusztításában vesznek-e részt, amelyen keletkeznek. A kompetíciós vizsgálatok tehát egy összetett folyamat egészére vonatkozó eredményt ígérnek viszonylag egyszerű technikai feltételek mellett.

Eddigi kutatásaink két, lényegében egymással ellentétes irányban folytak. Az egyik a természetesebb: ugyanúgy mint az egy-targetes esetben, meg kell találni a módját annak, hogy a mérések eredményeit megjósoljuk a bemért mennyiségek és a folyamat ideje alapján. Elvileg tehát ugyanazokat a módszereket alkalmazhatjuk itt is, amiket az egy-targetes esetre kifejlesztettünk. Mivel azonban minket elsősorban a gátlás és stimulálás kimutatása érdekelt, túlnyomóan a harmadik módszert alkalmaztuk, és azt vizsgáltuk, hogyan építhető be gátlás és stimulálás a folyamatot leíró differenciálegyenletekbe.

Azt tapasztaltuk, hogy a stimulálást egyszerűbb modellezni. Ennek legegyszerűbb módja, ha az effektor csak bizonyos előkészítés után pusztít, és az aktivált effektor hatása nem szelektív:

$$E' = -(f_A A + f_B B)E + (k_A A + k_B B)F,$$

$$F' = (f_A A + f_B B)E - (k_A A + k_B B),$$

$$A' = -k_A A F,$$

$$B' = -k_B B F.$$

Nem célunk itt a technikai részletekre kitérni, csak megjegyezzük, hogy amikor a munkába belekezdünk, egyáltalán nem volt világos előttünk, lehetséges-e az adott számítástechnikai feltételek mellett ilyen modelleket a mérési eredményekre illeszteni. Ma már elmondhatjuk, hogy a dolog ugyan lehetséges, de egyrészt nagyon időigényes, másrészt az eredmények meglehetősen megbízhatatlanok. Ez a két tapasztalás szorosan összefügg: ahhoz, hogy reális időn belül eredményt kapjunk, sok egyszerűsítésre, közelítésre, elhanyagolásra van szükség. Ezek együttes hatásának lemerése azonban nehezebb feladat, mint az eredeti illesztés. Egyszerűen annyit mondhatunk itt a tapasztalatainkról, hogy nem tartjuk a végeredményeinket

rossznak vagy megbízhatatlannak, csak sok nem várt nehézség jelentkezett munka közben.

Amint látható, a szokásos tömeghatás dinamikával mindkét target aktivizálja az effektort egymástól függetlenül, és az aktivizált effektor ismét egymástól függetlenül pusztítja őket, miközben inaktivizálódik. Itt tehát a serkentés csupán abban áll, hogy a nagyobb aktiváló erejű target mellett a másik jobban pusztul, persze az első kárára.

Más jellegű serkentés érhető el, ha az aktivizálás két lépésben történik. Ezt tükrözi a következő modell, amelybe sok más módosítást is beépítettünk.

$$\begin{aligned}E' &= -((r_A + f_A)A + (r_B + f_B)B)E, \\F' &= (f_A A + f_B B)E - (g_A A + g_B B)F, \\G' &= (g_A + g_B B)F - (k_A A + k_B B)G, \\A' &= -k_A AG, \\B' &= -k_B BG.\end{aligned}$$

Ha itt f_A és g_B kicsi, f_B , g_A nagy, a kompetitív helyzetben sokkal nagyobb, pontosabban mondva gyorsabb pusztulás tapasztalható, mint az egy-targetes ágakon. Ez tehát minőségileg más serkentés, mint ami az előző modellben van. Hasonló eszközökkel gátlást nem tudunk elérni, ezért szerepel a modellben az effektor direkt elpusztítása, ami a gyakorlatban úgy képzelhető el, hogy bizonyos feltételek mellett úgy kötődik meg az effektor a membránon, hogy ott csapdába esik, nem pusztít, de nem is tud kiszabadulni.

Tekintve, hogy a laboratórium csak a pusztulás mértékét tudja megmérni, nem elhanyagolható kérdés a fenti modellező tevékenységben az identifikálhatóság. Kezelhetjük persze a dolgot kezdetben pragmatikusan: az elsődleges cél a mérési eredmények jó, a mérési hibán belüli közelítése. Mi ezt még nem értük el, a mérési hiba a relatív pusztulásban 3-4%, a legjobb közelítéseink 5-6% körüliek, de sokszor 8-10%-nál nem lehet jobbat elérni egy adott modellben. Ebben a pragmatikus világban önálló feladatként jelentkezik annak eldöntése, van-e egyáltalán gátlás vagy stimulálás. Ezt képes megválaszolni a következő modell.

$$\begin{aligned}E' &= -(k_A + k_B B)E, \\A' &= -k_A AE, \\B' &= -k_B BE,\end{aligned}$$

ahol

$$\begin{aligned}k_A &= \exp(p_A + q_B B), \\k_B &= \exp(p_B + q_A A),\end{aligned}$$

Itt a q paraméterek előjele dönti el, hogy milyen irányú a kölcsönhatás a két target között. Nem fogalmaztuk meg, milyen feltételek mellett keresünk mi gátlást

és stimulálást leíró, tartalmazó modelleket, de e feltételek lényege a target-effektor kapcsolat: csak ez létezhet, és a kétféle target közvetlenül nem befolyásolhatja egymást. Tehát végül is felrúgtuk kimondatlan szabályainkat a cél érdekében, hiszen harmadik modellünkben az egyik target meghatározza a másik reakciósebességét. Így vált lehetővé objektíve lemérni, van-e kölcsönhatás köztük. Az eredmény természetesen csak viszonylagos lehet: az adott modell keretei között mondhatunk csupán véleményt.

A fenti modell már csak egyszerűsége folytán is sok kimondatlan feltételt tartalmaz, amelyeken megpróbálhatunk enyhíteni. Ezek közé tartozik az effektor és target reakciók szimmetrikus volta, ami eddig szintén része volt modelljeink szintaxisának. Ezen túlmenően arra is lehetőséget adott a modell, hogy a reakciósebesség függvényalakját meghatározzuk. Így jutottunk az

$$E' = -\alpha ET^{1+\kappa},$$

$$T' = -\alpha\beta ET^2$$

egy-targetes dinamikára, amelynek van explicit megoldása, viszont nem világos, hogyan lehet kiterjeszteni két targetre. Ismét egy hallgatólágos feltételezés lepleződik itt le: csak bizonyos egy-targetes modellek általánosíthatóak természetes módon több targetre. Ezért átmenetileg egy mechanikus kiterjesztést alkalmaztunk, amely viszont egészen általánosan elmondható.

Fokozatosan vált világossá a munka során a következő kérdés. Van-e a gátlásnak és stimulálásnak abszolút kritériuma, definíciója? Milyen feltételek mellett és milyen értelemben mondhatjuk azt egy adott folyamatra, hogy a targetek között interakció van? Megfordítva a kérdést: tegyük fel, hogy az egy-targetes dinamika tökéletesen ismert, vagyis adott a

$$P(T, E, A, I)$$

függvény, mely a target, effektor, antitest és idő függvényében megadja a pusztult target mennyiségét, és tartalmaz bizonyos, a targetre jellemző paramétereket. Egyáltalán, hogyan lehet egy ilyen függvény segítségével meghatározni azt a két-targetes modellt, amelyben nincs interakció? Van-e erre a célra automatikus eljárás? Mi nem ismerjük erre a kérdésre a választ:

- nem tudjuk, ez a kérdés felmerült-e már az irodalomban, ha igen, milyen válasz született eddig;
- nem tudjuk, egyáltalán értelmes-e a kérdésünk, és megválaszolható-e;
- szoktunk választ adni a kérdésre, ezekkel azonban nem vagyunk megelégedve, és válaszaink az idő során megváltoznak.

Most a következő az elképzelésünk. Meg kell határozni két pozitív állandót, mondjuk p -t és q -t, és a két targetes kísérlet pusztulásait a

$$pP_A\left(\frac{T_A}{p}, A_A, E, I\right), \quad qP_B\left(\frac{T_B}{q}, A_B, E, I\right)$$

mennyiségeknek kell feltételezni. Ha itt $p+q = 1$, akkor nincs kölcsönhatás, $p+q > 1$ esetén stimuláció, $p+q < 1$ mellett gátlás lép fel.

Első pillanatban talán meggyőzőnek hangzik az elképzelésünk. Pedig vannak hiányosságai: nem képes figyelembe venni az egyes folyamatok dinamikáját, mechanikusan eleve kettéosztja a két targetre ható effektort, és annak függvényében határozza meg a pusztulást.

Ezen a ponton mondunk búcsút ennek a témának, amellyel a többen részletebben foglalkoztunk, de hát végül is a legáltalánosabb kérdést is konkrét részleteire kell lebontani.

A rákos betegek gyógyítása általában hosszadalmas, több évig tartó folyamat, időszakonkénti intenzívebb kezeléssel. Ezért a gyógyítás nem választható el a betegek követésétől, és ez az az utóbbi sosem pusztán passzív megfigyelést jelent, hanem folytonos visszacsatolást magára a gyógyításra egyrészt konkrétan az egyes beteg esetében, másrészt általánosan a betegség egyes típusaiból álló csoportokra, új kezeléseket vezetnek be és régieket szüntetnek meg.

A betegek sorsának követése

Legyenek X és Y független, F illetve G eloszlású valószínűségi változók, legyen $i = 1, 2, \dots, n$ mellett (X_i, Y_i) az (X, Y) párra végzett független, egyforma eloszlású minta, és tegyük fel, hogy helyette nekünk csupán a

$$Z_i = \min(X_i, Y_i)$$

$$\Delta_i = \begin{cases} 1 & \text{ha } X_i \leq Y_i \\ 0 & \text{ha } X_i > Y_i \end{cases}$$

megfigyelések állnak a rendelkezésünkre. Kérdés, becsülhető-e F , és ha igen, hogyan.

Ez ma már klasszikus feladata a statisztikának, és egyben a betegkövetés alapfeladata is. Az utóbbi esetben X a vizsgált túlélés, Y az egyéb halálokok. Aszimptotikusan leeffektívebb becslést az ún. maximum likelihood módszer ad, amely az ismeretlen F eloszlást diszkrétnek tételezi fel és az eloszlás lehetséges értékei a Z_i mintaelemek. A kérdés csak az, hogy mekkora legyen a mintapontokba koncentrált tömeg.

Jelöljük a rendezett minta elemeit U_i -kel, és a megfelelő indikátor változót V_i -vel, és legyen

$$\lambda_i = P(X > U_i \mid X \geq U_i).$$

Akkor a maximálandó függvény

$$\prod_{k=1}^n (\prod_{i=1}^{k-1} \lambda_i - \prod_{i=1}^k \lambda_i)^{V_k} \left(\prod_{i=1}^k \lambda_i \right)^{1-V_k} =$$

$$= \prod_{k=1}^n (\prod_{i=1}^k \lambda_i) \left(\frac{1}{\lambda_k} - 1 \right)^{V_k} = \prod_{k=1}^n \lambda_k^{n+1-k} \left(\frac{1}{\lambda_k} - 1 \right)^{V_k}.$$

Itt az egyes tagoknak egymástól függetlenül a

$$\lambda_k = \begin{cases} (n-k)/(n-k+1) & \text{ha } V_k = 1 \\ 1 & \text{ha } V_k = 0 \end{cases}$$

veszi fel a maximumát, így a becslés szerint

$$P(X > U_k) = \prod_{i=1}^k \left(\frac{n-k}{n-k+1} \right)^{V_k}, \quad k = 1, 2, \dots, n.$$

Szavakban ez a becslés úgy fogalmazható meg, hogy balról jobbra göngyöltjük a keresett tömeget, és az egyes mintapontokba pontosan akkor teszünk le belőle egyáltalán valamennyit, ha az a pont valódi minta, és ebben az esetben a hátralevő mintapontok száma szerint arányosan erre a pontra jutó tömeget kell letennünk. Előfordulhat, hogy az utolsó mintapont után marad a kezünkben ki nem osztott tömeg, annak a helyét a minta nem határozza meg, csak annyit mondhatunk róla, hogy nagyobb a legnagyobb mintaelemnél.

E becslés aszimptotikus viselkedésének a vizsgálatában mi is részt vettünk, elméleti szempontból is jelentős a feladat, mert sok hasonló kérdés megoldásának vált kiindulópontjává. Ezek közül a legfontosabb a kísérő változók hatásának a figyelembe vétele. Mi módon jelenik meg az élettartam alakulásában a sokféle külső tényező? Vagy az operációt megelőző körülmény, önmagának a szóbanforgó betegségnek a konkrét paraméterei, a beteg életkora, a kezelés egyes jellemzői? Ha nem találjuk meg a módját, hogy közös nevezőre hozzuk a sok egyéni sorsot, tudományunk a kórtörténetek sorozatára redukálódik, és a sok különböző előjelű hatás között tétovázva nem találjuk meg a gyógyítás útját.

Hasonló feladat az iparban is előfordul: műszaki berendezések élettartamát nem vizsgálhatjuk azok rendeltetésszerű használatával, hiszen erre nincsen idő. Túl kell terhelniük a berendezéseket, és a megnövekedett, felszított igénybevétel alatti megrövidült élettartamból kell következtetnünk a valódi élettartamra. Cox vizsgálta először azt a hipotézist, hogy a túlélési függvény kitevőjében jelentkezik a hatás:

$$P(X > t) = (G_0(t))^{\exp \sum_{i=1}^k c_i Y_i},$$

ahol X a vizsgált élettartam, k a kísérő változók száma, Y_1, Y_2, \dots, Y_k a kísérő változók, és c_i a regressziós együtthatók.

A gyakorlatban ez a modell a bevezetőben tárgyalt, ún. cenzoráló hatással együtt jelentkezik. Így már nem adható explicit formula a maximum likelihood becslésre, de annak kiszámolására effektív iterációk léteznek.

A $G(t) = P(X > t)$ túlélési függvényt logaritmikus deriváltjának, a

$$\rho(t) = \frac{f(t)}{G(t)} = -(\log G(t))'$$

függvénynek az alakjával jellemezhetjük, ez a meghalási intenzitás, hiszen közelítőleg $\rho(t)\Delta t$ a feltételes valószínűsége kis Δt mellett annak, hogy t és $t + \Delta t$ között hal meg az X élettartamú egyén feltéve, hogy a t időben még él. Cox modellje szerint külső hatások a meghalási intenzitást szorozzák, mintegy az idő múlását mérő óra skáláját változtatják meg. Független halálokok esetén a megfelelő ρ függvény összeadódik, ami a túlélési függvényben egyszerű szorzó faktort jelent. A Cox regresszió tehát kifejezetten kölcsönhatást tükröz, nem párhuzamos történéseket. Nem arról van szó, hogy a dohányzás önmagában elindít egy folyamatot, ami a nem-dohányosokra nem is létezik, csak a dohányos órája gyorsabban ketyeg. Amit mondunk, csak példa: nem a mi dolgunk itt a dohányzás hatásáról beszélni, csak azt akarjuk szemléltetni, hogy mit jelent önmagában a Cox modell. Amint látható, nem-paraméteres abban az értelemben, hogy tetszőleges $G(t)$ túlélési függvény mellett alkalmazható, a statisztikus feladata a mérési, megfigyelési eredmények alapján a $G_0(t)$ függvényt és a c_i regressziós együtthatókat meghatározni, természetesen e becslések megbízhatóságának a meghatározásával együtt. Ezért, mármint a modell nem-paraméteres jellege miatt nem fordíthatóak le hétköznapi nyelvre a c_i regressziós együtthatók automatikusan. A példánál maradva nem mondhatjuk abszolút értelemben azt, hogy napi 1 cigarettával több 2 hónappal rövidíti meg az életet. Minden, a várható élettartamban jelentkező hatás a háttérben álló $G_0(t)$ túlélési függvénytől függ. Szemléltetésül persze feltételezhetjük, hogy ez legyen mondjuk exponenciális függvény, vagy maga a becsült érték — a fontos csak az, hogy ezek az értelmezések a feltevésünktől függenek, csak akkor lesznek az eredményeink másokéval összehasonlíthatóak, ha világosan megmondjuk, mi melyik értelmezést használjuk.

Elméletileg sokkal bonyolultabb feladat a többdimenziós élettartamok vizsgálata. Erre akkor van szükség, ha például a szervek élettartamát külön vizsgáljuk, mondjuk a primér daganat helye és az áttételek jelentkezése szerint. Ebben az esetben X és Y többdimenziós vektorok, Z a koordinátánkénti minimumok vektora és Δ a megfelelő indikátor vektor. A maximum likelihood becslés nem adható meg explicit alakban, és meghatározása a következő feladatra vezet. Adott vektorokhoz, amelyek koordinátái pozitívak, keresendő az a vektor, amelynek a koordinátái ugyancsak pozitívak, a koordináták összege 1, és amelynek az adott vektorokkal képzett skaláris szorzatainak a szorzata maximális. Ez a feladat is megoldható az ún. EM algoritmussal, amelynek konvergenciájára Csiszár Imre adott bizonyítást. A tapasztalat azt mutatja, hogy az algoritmus nem elég gyors, és más, a nem lineáris programozás területén kifejlesztett algoritmusok effektívebbek nála. Mivel konkáv célfüggvényt kell lineáris kényszerfeltételek mellett maximálni, Bregman algoritmus is alkalmazható.

Ezeknél a feladatoknál nem lényeges feltétel az, hogy a cenzoráló változó is független és egyforma eloszlású legyen. Egészen általánosan úgy fogalmazható meg a feladat, hogy egy ismeretlen valószínűségi mértékre vonatkozó független, egyforma eloszlású minta helyett csupán bizonyos halmazokat kapunk, amelyekben a minta elemei rendre benne vannak és amelyek a minta elemeitől valamilyen előre elhatározott módon függenek. A megoldatlan kérdés az, hogyan lehet jellemezni azokat a megfigyelési módokat, amelyek mellett egyáltalán becsülhető az ismeretlen mérték,

és mekkora ennek a becslésnek a megbízhatósága.

Technikailag a betegek követése adatbankok létesítését jelenti. Igen megnehezíti a munkát az állandó változás. Egy-egy felmérés megtervezése, az adatok összegyűjtése, azok kiértékelése olyan lassú, hogy közben az esetek többségében az egész munka érvényét veszti, megváltozik az együttműködésben részt vevő szakmai közösség összetétele, érdeklődése, megváltozik a kezelés, újabb gyógyszerek kerülnek forgalomba. A betegek követése, állapotuk leírása egyre több részletre terjed ki, amik természetes módon a korábbi tervekben még nem szerepeltek. Ezért a statisztikus nagyon nehezen talál a klasszikus módszerekkel kiértékelhető anyagot. Nem elhanyagolható nehézség volt az elmúlt években a számítástechnikai háttér folyamatos változása. Elvileg persze semmi nem gátolhatja meg azt, hogy pontosan azt a folyamatot vizsgáljuk, ami az adott esetben a betegek sorsából és az őket követő rendszer változásaiból kialakul, de az ilyen bonyolult esetekben nagyon nehéz megtartani a statisztikának az objektivitását, kiszűrni a vizsgált folyamat és megfigyelése közti interakció torzító hatását. Azt például, hogy a ritka tüneteket egy idő után már nem is regisztrálják, a gyakoriakat pedig a vizsgálat során egyre részletesebben írják le. Nagyon nehéz figyelembe venni a beteg és az orvos szubjektív döntéseit, egy jó orvos nem tudja világosan megmondani, mikor mit miért csinál. Sokszor csupán tájékozódni szeretne. Ha a kiértékelés kedvéért megkötjük a kezét, csökken a munkája effektivitása, viszont ha nem követi az eredeti terveket, az összegyűjtött anyag válik értékelhetetlenné.

Kidolgoztunk egy általános eljárást, amelynek segítségével tetszőleges változók alapján lehet diszkriminálni, vagyis előre adott csoportokhoz osztani a mintát. Minden egyes magyarázó változó minden egyes értéke kap mindegyik csoporthoz egy számot, és az egyes eseteket a csoportok számai összegének a maximuma alapján minősítjük. A módszer paramétereinek a meghatározása eléggé munkaigényes, de nem lehetetlen. Egy nagy iteráció a lényege, amelyik parciálisan minden egyes ismeretlent meghatároz feltéve, hogy a többinek annyi az értéke, ahol éppen az iteráció tart. Látható, hogy az algoritmus szeletelhető, mint a salámi, a magyarázó változók nagyobb csoportja együtt mozgatható benne, vagy már feldolgozott változókhoz újabbak csatlakoztathatóak. Figyelemre méltó a túlillesztés veszélye, különösen ha sok magyarázó változónk van, ami a szokásos osztott-minta technikával csökkenthető: a minta véletlenszerűen kiválasztott része alapján becsülünk, más része alapján ellenőrizzük a kapott paramétereket.

Ennek a módszernek nagyon egyszerű a logikai struktúrája, hiszen egyenrangúak benne a magyarázó változók. Sokszor effektívebbek a hierarchikus módszerek: néhány változó alapján durva, előzetes osztályozást alakítunk ki, és az egyes osztályok diszkriminációjára az osztályra jellemző változókat használjuk. Ezek az általános, a konkrét esetre vonatkozó előzetes információkat nem használó módszerek igen nagy mintákat igényelnek. És a munka kezdetén sokszor az a kellemetlen érzése van minden résztvevőnek, mint amikor süketek beszélgetnek. Hiszen a statisztikus első eredményeit az orvos vagy ismeri, amikor is feleslegesnek érzi a számításokat, vagy nem ismeri, amikor megbízhatatlannak tartja azokat. Szerencsésebb az együttmű-

kódés, ha a statisztikus képes beépíteni a vizsgálatba a meglevő ismereteket. A legjobb ezt már az adatgyűjtés tervezésekor megtenni azzal együtt, hogy világosan megfogalmazzuk a konkrét munka célját.

Zárszó

Matematikai oka van-e a ráknak? Szükségszerű-e, hogy a sejtek másolási láncában meglegyen a lehetősége annak a változásnak is, amely agresszívebb, a saját szempontjából életképebb példányokat hoz létre? Kapcsolatban áll-e ennek a valószínűsége a szervezet bonyolultságával? Van-e kapcsolat a rák kialakulása és az életmód között, van-e szerepe az életmódnak a gyógyulásban?

Szerényebben fogalmazva: van-e a matematikának olyan ága, amelyik speciálisan a rákkutatáshoz kell? Mi nem ezt kerestük. A saját munkáját végző klinika támogatására megpróbáltuk megadni mindazt a matematikai segítséget, ami a meglevő és várható tevékenységéhez nélkülözhetetlen.

A matematikai módszerek zöme tíz-húsz éve alakult ki, és ma már rutinnak számít. Egyrészt az irodalom összegyűjtése, másrészt a szakemberek segítése lehet a feladatunk, beleértve a számítógépek programrendszerait. Ilyen jellegű például a tomográfia, amiről itt most nem beszéltünk, de világos matematikai vonatkozásai vannak. Mi változatlanul szerencsétlen, szükséges rossz dolognak tartjuk a kis, személyi számítógépek itthoni egyoldalú terjedését, szerintünk a klinikának nagyon sürgősen egy valódi, nagy kapacitású központi gépre, és körülötte működő csoportra van szüksége, amelyik természetes módon kapcsolatban áll a saját személyi gépekkel dolgozó kisebb csoportokkal. De ez részletkérdés.

Ami a kutatást illeti, az nyilvánvalóan csak nemzetközi lehet. Mi elsősorban a DNS és fehérje-láncok feltérképezésétől várjuk az élet kialakulásának, és a rák fejlődésének a megértését. Mennyiségét tekintve is hatalmas munka ez, szerteágazó, ami a törzsfejlődést illeti, és matematikus, vegyész, biológus és orvos együttműködését igényli.

Történelmi léptékben mérve az egyedek és fajok láncolata alakította ki a létező fehérjéket, azt a laboratóriumot, amit a sejtek belső élete jelent. Az egyedek életében ez a sejtek újratermelése során felgyorsulva megismétlődik, természetesen sokkal kisebb változékonysággal. Ma már tudjuk, hogy az immunrendszer teljes fehérjekészlete nincs a DNS-ben kódolva. Csak a mozaikok, és a mód, hogy ezekből mi módon kell a hosszú láncokat felépíteni. Mondhatjuk, hogy nincs kilistázva a DNS-ben a fehérje, csak be van programozva. De ezt a programnyelvet csak most kezdjük megtanulni. Lehet a rák programhiba. Nem abban az értelemben, hogy megváltozik a DNS-lánc, hanem úgy, hogy e változás eredményeképpen továbbra is működőképes program jön létre, de az új program már nem azt írja le, amit az eredeti. Akik gépen dolgoznak, tudhatják, milyen nehéz az ilyen hibákat lokalizálni. Ha nem működik a program, a futás egy ponton megáll, a hiba könnyen megtalálható. De ha minden inputra van a programnak outputja, csak azok közül néhány

téves, ezt a hibát még észrevenni is nagyon nehéz.

Tehát meg kell ismerni a sejtek osztódásában beálló változások természetét, és meg kell ismerni a sejt-laboratórium működését. Ennek a hosszú folyamatnak néhány részletéről beszéltünk a fentiekben.

Önmagában nagy kérdés, milyen mértékben függenek ezek a folyamatok a véletlentől. Ha ismernénk egy adott beteg minden egyes sejtjének tökéletesen a tartalmát, akkor meg tudnánk-e mondani, hogy a különböző lehetséges kezelések hatására mi fog történni? A beteg-követések ma még szükségképpen felületesek abban az értelemben, hogy a figyelembe nem vett hatások együttesen nagyobbak lehetnek a véletlen hatásánál.

Hatalmas irodalma van a csodálatos gyógyulásoknak. Természetes dolog: akik meggyógyulnak, ezt örömmel elmesélik. Mi módon lehetséges kiszűrni a szelekció torzító hatását? A matematika önmagában ezt nyilván nem mondhatja meg. De az sem valószínű, hogy matematika nélkül lehet a kérdésre válaszolni.

(Beérkezett: 1990. december 20.)

TUSNÁDY GÁBOR
MTA MATEMATIKAI KUTATÓ INTÉZETE
1053 BUDAPEST, REÁLTANODA U. 13-15

MATHEMATICAL METHODS OF CANCER RESEARCH

G. TUSNÁDY

The paper contains three parts:

1. Understanding and description of processes leading to cancer: a possible source for this aim may be the investigation of non ergodic sequences of stochastic automatae.
2. Diagnostic methods in histology: the proportion of different lines in evolution and their growth factor may be estimated by algorithms developed and used for decomposition of mixtures.
3. Follow-up studies: we used Kaplan-Meier estimator and Cox regression for one and multidimensional survival distributions.

A külföldi szakirodalomból

KÁOSZ ÉS A VÁLASZTÁSOK ELMÉLETE*

D. G. SAARI

Ewanston

1. Bevezető

A „káosz” fogalma egy meglehetősen új kutatás eredménye, de máris fontos, lényegbevágó szerepe van a dinamikus rendszerek tanulmányozásában. Ezen matematikai eljárások felhasználásával kezdjük megérteni, hogyan és miért mutathatnak determinisztikus rendszerek (deterministic systems) véletlenszerűnek tűnő jelenségeket (lásd pl. Devaney, 1986.). Például ezek az új megközelítési módok vittek minket közelebb a numerikus analízis, valamint a fizikai és a biológiai rendszerek megértéséhez. Újabban, amint az várható volt, a káosz jelenségét felfedezték a társadalomtudományokra kifejlesztett dinamikus rendszerekben is. Az, hogy a káosz jelenségét egyre gyakrabban fedezik fel a tudományos kutatások legkülönbözőbb területein alkalmazott dinamikus rendszerekben, nem meglepő. Ami váratlannak tűnhet, az az, hogy a korábban csak „statikus” összefüggésben vizsgált problémák szintén analizálhatók ily módon. A statikus problémákat egy kvázidinamikus keretbe helyezve megérthetjük, összehasonlíthatjuk és általánosíthatjuk a döntéselméletben, a valószínűségszámításban, a diszkrét programozásban és a tudomány számos más területén előforduló paradoxonokat. Ily módon sok új és váratlan látószöveg ellentmondásra derül fény. Mitöbb, ha dinamikus rendszerek fejlődését vesszük iránymutatónak, egy új program alakul ki e témák vizsgálatára. E cikk célja az, hogy ezt egy konkrét példán mutassa be. Egy, a választásokkal kapcsolatos régi döntéselméleti probléma új megoldását fogom leírni.

A következő példával szeretnék illusztrálni egy olyan problémát, amellyel a választások során szembekerülünk. Tegyük fel, hogy van kilenc választónk, akiknek három jelölt (A , B , és C) között kell választania. Tegyük fel továbbá, hogy négyen $A > C > B$ sorrendet, hárman $B > C > A$ -t, és az utolsó kettő $C > B > A$ sorrendet állítja fel a jelöltek között. Ha a standard választási rendszert használjuk, amelyben minden egyes szavazó a saját első helyezettjére voksol, akkor a választás

*A fordítás D. G. SAARI, Chaos and the theory of elections, in: Kurzhanski, A. B. and Sigmund, K. (eds) *Dynamical Systems* (Lecture Notes on Economical Mathematical Systems, vol. 287) Springer, Berlin–Heidelberg–New York, 1987, pp. 179–188 dolgozatából a szerző és a kiadó engedélyével készült.

eredménye $A > B > C$ a $4 : 3 : 2$ arányban. Ez a sorrend azonban ellentmondásban van azzal a ténnyel, hogy a szavazók többsége (a kilencből öt) a második helyezett B jelöltet jobban kedveli az első helyezett A jelöltnél. Még szembetűnőbb az, hogy a szavazók kétharmada előbbre helyezi az utolsó helyezett C -t B -vel szemben és a többség előnyben részesíti C -t A -val szemben. Ezek páronkénti, tranzitív sorrendek (relációk) a $C > B > A$ sorrendet határozzák meg, amely pontosan ellentettje a választási sorrendnek. Még durvábban fogalmazva, C a többségi jelölt (aki bármely páronkénti összevetésben a többségi szavazatot bírja), de a választáson utolsó helyezett. A legnépszerűtlenebb jelölt A (aki elveszti a választást valamennyi páros összevetésben) viszont élre kerül.

Az amerikai és a francia forradalom idején az Atlanti óceán mindkét partján fejlődtek azok az eljárások, melyek egyéni preferenciák alapján csoport rangsort állítottak fel. Részben ezektől a vitáktól sarkallva fedezték fel a fentiekhez hasonló példákat. Hamarosan felismerték, hogy ezek a paradoxonok azért fordulnak elő, mert minden egyes szavazó csak a saját első helyezett jelöltjére szavazhat. BORDA nagy hatású munkájában (1781) egy egyszerű megoldást indítványozott, amelyben minden egyes szavazó megadja az N számú jelölt sorrendjét. Ezután a szavazatok kiértékelése (tallying process) során egy szavazó j -edik helyezett jelöltje $(N - j + 1)$ pontot kap ($j = 1, \dots, N$). Egy jelölt végleges helyezését az így kiszámított pontok száma határozza meg (konkrétan, Borda javaslata $N = 3$ -ra szolt). Például a fenti esetben, Borda $C > B = A$ sorrendje a $20 : 17 : 17$ pontozással közelebb van a látzólag helyes $C > B > A$ sorrendhez (amely a páronkénti összevetésből jön).

Ez ésszerű megoldásnak tűnik, de mégsem mentes a fogyatékoktól. Borda módszerének kritikáját Laplace és mások vetették fel, három fő pont köré csoportosítva azt. Ezekből kettőt vitatunk itt meg. Az első a szavazatok kiértékelése során alkalmazott súlyokkal foglalkozik: mivel igazolható, hogy éppen ezt a pontozást alkalmazzuk? A $(3, 2, 1)$ helyett miért nem használjuk a $(4, 2, 1)$ -et? A felhasználható súlyoknak meg kell felelniük bizonyos nyilvánvaló feltételeknek. Bármely súly kiválasztását tekinthetjük úgy, mint egy szavazási vektor $W = (\omega_1, \dots, \omega_N)$ definiálását, amely eleget tesz az $\omega_j \geq \omega_{j+1}$ ($j = 1, \dots, N - 1$) monotonitási feltételeknek.

A célból, hogy a pontozás (kiértékelés) különbséget tegyen a jelöltek között, kiköthetjük, hogy $\omega_1 > \omega_N$. Az egyszerűség kedvéért, s mert ez mindig teljesül a gyakorlatban, azt is feltesszük, hogy valamennyi ω_j racionális szám. A kiértékelésben ω_j pontot kap egy szavazó j . helyezett jelöltje, amely ilyen módon egy súlyozott szavazási rendszert határoz meg. (Például a standard szavazási rendszert a $(1, 0, \dots, 0)$ vektor határozza meg, míg Borda rendszerét az $(N, N - 1, \dots, 1)$ súlyok definiálják.) A kérdés, melyet Borda kritikussai vetettek fel, tekinthető úgy, mint annak a kritériumnak a keresése, amely igazolja egy konkrét W kiválasztását a végtelen sok lehetőség közül.

A második kritika az volt, hogy Borda módszere nem mindig a „többségi” jelöltet választja ki. Ennek illusztrálására tekintsük az alábbi 11 szavazó példáját, ahol hatan a $B > A > C$, négyen az $A > C > B$, míg egy szavazó a $C > B > A$ sorrendet jelöli meg. A Borda módszer az $A > B > C$ választási sorrendet adja $25 : 24 : 17$ arányú pontozással. Ugyanakkor a 2. helyezett B jelölt nyeri a többségi szavazatot,

akár A -val, akár C -vel hasonlítjuk össze. Tehát, bár B a többségi jelölt, mégsem áll az élen Borda számítása szerint. Mindazonáltal ebben az esetben a standard választással adódó $B > A > C$ sorrend megegyezik a páronkénti összevetésből adódó természetes sorrenddel.

Most már többet tudunk a szavazási rendszerekről és a választások elméletéről. ARROW találta meg (1963) a legfontosabb eredmények egyikét: Bebizonyította, hogy nincs olyan eljárás, amely kielégítene bizonyos egyszerű, ésszerű axiómákat. Eredményének közvetlen következménye az, hogy a fenti példák nem elszigetelt jelenségek. Akárhogy választunk ki egy szavazási vektort, egy szavazócsoporthoz két jelöltre kialakított rangsora nem szükségképpen egyezik meg azzal, ahogy ugyanaz a csoport rangsorolja a párost három vagy több jelölt esetén. Amíg ez, és számos más eredmény rávilágított azokra a pontokra, melyeket Borda kritikussai vetettek fel, az alapvető kérdések mostanáig megválaszolatlanok maradtak. A megoldáshoz, amelyet a következőkben főbb vonalaiban leírnunk, a káoszról és a szimbolikus dinamikából kölcsönözünk ötleteket.

2. Szimbolikus dinamika

A szavazással kapcsolatos vizsgálatokban jelentős szerepet játszanak az iterációs eljárások dinamikájából vett gondolatok. Tekintsük az alábbi iterációs eljárást

$$(1) \quad x_{j+1} = f(x_j) \quad j = 0, 1, 2, \dots$$

Ha $p = x_0$ a kezdeti feltétel, akkor a cél az, hogy megismerjük a pályagörbét, azaz

$$(2) \quad \{f^{(k)}(p)\} \quad k = 0, 1, 2, \dots$$

sorozatot, ahol $f^{(k)}$ az f -nek az önmagával vett k -adik kompozíciója. A káosz leírásának egy fontos eszköze a szimbolikus dinamika. Ebben az eljárásban az állapottér (statespace) bizonyos tartományai fel vannak osztva. Minden egyes elkülönített halmaz egyértelműen meg van jelölve egy jellel, például az ábécé egy betűjével. Ekkor, egy meghatározott kezdeti feltétel esetén a (2) egyenlet által definiált pályagörbe a szimbólumok egy sorozatát határozza meg, ahol a sorozat $(k+1)$ -edik eleme az $f^{(k)}(p)$ -t tartalmazó halmazhoz tartozó jel. Egy ilyen sorozatot „szónak” nevezünk, és az összes lehetséges szavak halmazát „szótárnak” hívjuk. Ha tudjuk, hogy milyen szavakat tartalmaz egy szótár, akkor mélyebben megértjük a rendszer dinamikáját. Az a szélsőséges eset, amikor minden lehetséges szó előfordul: a káosz egy formája.

A szimbolikus dinamika ezen megközelítését használjuk arra, hogy megválaszoljuk Borda kritikussait és szavazási rendszereket analizáljunk. Hogy ezt megtegyük, egy „dinamikus rendszert” terveztünk a választások számára. Az eljárást egy több függvényből álló halmaz $\{f_k\}$ határozza meg, amely kijelöli a választási eredményeket. A „szavazás pályagörbéjét” az

$$(3) \quad \{f_k(p)\}, \quad k = 1, \dots, T$$

alakban feltételezzük, ahol az index azonosítja a jelöltek egy rangsorolandó részalmazát.

A szavazás célja az, hogy a jelöltekből álló adott részalmazra megadja a csoport-sorrendeket.

N jelöltből álló $\{\alpha_1, \dots, \alpha_N\}$ halmaz esetén $2^N - (N + 1)$ db 2 vagy több jelöltet tartalmazó részalmaz van. Legyen F néhány részalmaz családja, azaz legyen $F = \{S_1, \dots, S_T\}$, ahol S_j egy két vagy több jelöltet tartalmazó részalmaz, ahol $S_j \neq S_k$, ha $j \neq k$.

Például a fenti példákban egy családot használtunk; ez a család a jelöltekből alkotott, összesen $N(N - 1)/2$ db párból és az összes $(N \text{ db})$ jelöltet tartalmazó halmazból áll. A többségi szavazásnak és a Borda-módszernek a kritikája azon alapul, hogy megmutatja, a választási eredmények ezen család eltérő részalmazait tekintve nem adnak azonos eredményt. Újabb kritikák a jelöltek más részalmazaira támaszkodva nyújthatnak eltérő választási eredményeket; az F családot a részalmazok kiválasztása határozza meg.

Legyen adva egy $F = \{S_1, S_2, \dots, S_T\}$ család. Az összes választási eredmény meghatározása céljából, legyen R_j az S_j , $j = 1, \dots, T$ részalmazokban lévő jelöltek valamennyi lineáris sorrendjének halmaza. (Például, ha $S_j = \{\alpha_1, \alpha_2\}$, akkor $R_j = \{\alpha_1 > \alpha_2, \alpha_1 = \alpha_2, \alpha_1 < \alpha_2\}$). Ezek a sorrendek játsszák a szimbólumok szerepét. Legyen $W_F = (W_1, \dots, W_T)$ T db szavazási vektor halmaza, ahol W_j -t az S_j -beli jelöltekre leadott szavazatok kiértékelésére használjuk, $j = 1, \dots, T$. Ezek a vektorok, amelyek meghatározzák a választási eljárást F minden egyes részalmazán, leírják a „dinamikát”, azaz a (3) egyenlet leképezéseit. Hogy a leírást teljessé tegyük, hadd emlékeztessék arra, hogy a szavazók p „profil”-ja egy lista, amelyen a jelöltek sorrendjét megadják. Ha P jelöli az összes lehetséges profil terét, akkor S_k -ra vonatkozó választási eredmény az

$$(4) \quad f_{kW} : P \rightarrow R_k$$

módon adott, ahol f_{kW} az a kézenfekvő leképezés, amelyet a W_k kiértékelő eljárás definiál. Ez a leképezés a következőképpen van meghatározva. Ha W_k és P adott, a választási eredmény kiszámolható. Az aktuális számszerű végeredmény helyett egy szimbólumot használunk, amely a jelöltek választási sorrendjét írja le. Ezt jelöli $f_{kW}(p)$. Így egy adott W_F és p esetén F valamennyi részalmazára vonatkozó választási eredményeket a következő formában kapjuk:

$$(5) \quad \{f_{kW}(p)\} \quad k = 1, \dots, T.$$

Ennek illusztrálására tekintsük a következő példát, amelyben $N = 4$, $F = \{\{\alpha_1, \alpha_2\}, \{\alpha_2, \alpha_3\}, \{\alpha_1, \alpha_2, \alpha_3\}, \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}\}$ és $W_F = (1, 0; 1, 0; 3, 2, 1; 1, 0, 0, 0)$. Ha W_F -et így választjuk, az azt jelenti, hogy az első 2 részalmazon a sorrendet a szavazattöbbség határozza meg, a 3. részalmazt Borda módszerrel értékeljük ki, és az utolsó részalmazon a standard szavazással állapítjuk meg a sorrendet. Tegyük fel, hogy a P profil a következőképpen fest: 5 szavazó az $\alpha_1 > \alpha_4 > \alpha_2 > \alpha_3$ sorrendet, 4 az $\alpha_2 > \alpha_3 > \alpha_4 > \alpha_1$ és másik 4 szavazó az $\alpha_3 > \alpha_4 > \alpha_2 > \alpha_1$ sorrendet

adja meg. Ekkor $f_{jW}(p)$ az F család S_j ($j = 1, 2, 3, 4$) részhalmazainak választási sorrendje, ahol a választások kiértékelése az adott W_F választási vektor alapján történik. Az adott p -re ennek az eredménye a következő: $f_{1W}(p) = \alpha_2 > \alpha_1$ 8 : 5-ös pontozással; $f_{2W}(p) = \alpha_2 > \alpha_3$ 9 : 4-gyel; $f_{3W}(p) = \alpha_4 > \alpha_2 > \alpha_1$ 30 : 25 : 23-mal és $f_{4W}(p) = \alpha_1 > \alpha_2 = \alpha_3 > \alpha_4$ az 5 : 4 : 4 : 0 pontozási eredménnyel. Az (5) által definiált sorozat a következő

$$(6) \quad (\alpha_2 > \alpha_1, \alpha_2 > \alpha_3, \alpha_4 > \alpha_2 > \alpha_1, \alpha_1 > \alpha_2 = \alpha_3 > \alpha_4).$$

Egy adott F család esetén legyen D_F az $R_1 \times \dots \times R_T$ Descartes sorozat. D_F egy univerzális halmaz; D_F elemei a sorrendekből alkotott sorozatok, mind a T darab F -beli halmazhoz; és D_F tartalmaz minden lehetséges sorozatot. Ilyenformán D_F tartalmazza az F családra vonatkozó összes lehetséges és lehetetlen választási eredményt. Például, ha $F = \{\{\alpha_1, \alpha_2\}, \{\alpha_2, \alpha_3\}\}$, akkor D_F 9 sorozatot tartalmaz ($|D_F| = 9$). Ha $N = 3$ és F az összes lehetséges ($2^3 - 4$ db) részhalmazok halmaza, akkor $|D_F| = 351$; ha $N = 4$ és F az összes ($2^4 - 5$ db) részhalmazból álló család, akkor $|D_F| = 1686498489$.

Ahhoz, hogy egy W_F választási vektor hatását megvizsgáljuk, tudnunk kell az F családra vonatkozó összes lehetséges választási eredményt. Ez egy D_F -beli részhalmaz, amelyet úgy kapunk, hogy a W_F által adódó lehetséges választási eredményeket tekintjük.

$$(7) \quad D_{FW} = \{[f_{kW}(p)] \quad k = 1, \dots, T, \quad p \in P\}$$

Ha követjük a dinamikus rendszerek terminológiáját, akkor D_{FW} -t a W_F által definiált *szótárnak*, és D_{FW} -beli sorozatot *szónak* nevezhetjük. Például a (6) formulával megadott sorozat a kiválasztott W_F -hez tartozó szó D_{FW} -ben.

Az eredményeink megfogalmazásához W_F -re egy ekvivalenciarelációt szükséges megadni. Az első állításunk valamelyest hasonlít a topologikus ekvivalencia fogalmához.

Szeretnénk megtudni, hogy mikor ad W'_F és W''_F azonos pályagörbék, azaz bármilyen p választása esetén mikor fog 2 választási vektor mindig ugyanolyan választási eredményt adni. Ha ez teljesül, akkor W'_F -et és W''_F -t ekvivalensnek mondjuk. A következő állítás bizonyítása egyszerű (rutin) gyakorlat:

1. ÁLLÍTÁS. Tegyük fel, hogy adott egy F család. Két vektor, $W'_F = (W'_1, \dots, W'_T)$ és $W''_F = (W''_1, \dots, W''_T)$ akkor és csak akkor ekvivalens, ha bármely $j=1, \dots, T$ esetén a W'_j, W''_j és az E_S vektorok mindegyike egy megfelelő dimenziójú Euklideszi tér ugyanolyan kétdimenziós alterében vannak. Itt az E_S vektor dimenziója megegyezik $|S_j|$ -vel, azaz S_j elemszámával.

Mivel ez az azonosság fennáll, feltesszük, hogy ha S_j csak 2 jelöltet tartalmaz, akkor a hozzárendelt választási vektor az $(1, 0)$ vektor. Ez a vektor határozza meg a többségi szavazat kiszámításának módját. Hasonlóan könnyű megmutatni azt is, hogy ha a választás vektorban az egymást követő súlyok közötti különbség állandó,

akkor a vektor ekvivalens a Borda féle vektorral, azaz pl a $(45, 20, -5)$ ekvivalens $(3, 2, 1)$ -gyel. Ha W_F összes komponense ekvivalens $(1, 0)$ -val, vagy a Borda féle vektorral, akkor B_F -fel jelöljük. A B_F -hez tartozó szótárt pedig D_{FB} jelöli.

3. Választási eredmények

A Borda módszer kritikája alapján felvetődő kérdések matematikai megfogalmazásának egyik módszere lehet a D_{FW} és a D_{FB} szótárak összehasonlítása. Tegyük fel például, hogy az egyik szótár több szót tartalmaz, mint a másik. Ez azt jelenti, hogy a hozzátartozó választási eljárás F -re vonatkozólag több inkonzisztens választási eredményt enged meg, mint a másik. (Más szavakkal, mint a dinamikus rendszerek esetében, azt mondanánk, hogy minél nagyobb a szótár, a lehetséges választási eredmények annál kaotikusabbak). Másik oldalról pedig elmondható, hogy az az eljárás, amelyekhez kisebb szótár tartozik, könnyebben „megjósolható”, konzisztensebb eredményeket ad. Tehát adott F esetén, a D_{FB} , D_{FW} és a D_F közötti kapcsolatot fogjuk vizsgálni.

Könnyű olyan F családot konstruálni, ahol akármilyen W_F választás esetén $D_F = D_{FW}$. Ilyen például az $F = \{\{\alpha_1, \alpha_2, \alpha_3\}, \{\alpha_4, \alpha_5, \alpha_6\}\}$. Az a lényeg, hogy a két halmaznak ne legyen közös eleme, s így nincs mit összehasonlítani. Ez indokolja a következő definíció bevezetését.

1. Definíció. Azt mondjuk, hogy egy F család kielégíti a „bináris tartalmazási tulajdonságot” (a továbbiakban (b.t.t.))

(i) ha F tartalmaz legalább egy olyan halmazt, amelyben három, vagy több jelölt van,

és

(ii) ha S_j az F -nek egy olyan halmaza, amelyben három, vagy több jelölt van, akkor az S_j -beli jelöltek között van egy olyan pár, amelyik valamelyik másik F -beli halmazban is benne van.

Az egy halmazból álló család, a jelöltek összes $(N(N-1)/2)$ párját tartalmazó család, valamint a fenti két halmazos család nem tesz eleget (b.t.t.)-nek. Másrészt viszont a jelöltekből képzett összes párt és az $\{\alpha_1, \alpha_2, \alpha_3\}$ halmazt is tartalmazó család kielégíti (b.t.t.)-t. Mellesleg F bármely részcsaládja, amely tartalmazza az $\{\alpha_1, \alpha_2, \alpha_3\}$ és az $\{\alpha_1, \alpha_2\}$ halmazt, eleget tesz (b.t.t.)-nek. Az $F = \{\{\alpha_1, \alpha_2, \alpha_3\}, \{\alpha_4, \alpha_5, \alpha_6\}, \{\alpha_4, \alpha_6, \alpha_8\}\}$ nem elégíti ki (b.t.t.)-t, de az utolsó két halmazt tartalmazó részcsaládja igen. Az első eredményünk egy negatív eredmény.

1. TÉTEL (Saari, 1984, 1985 b). Tegyük fel, hogy $N \geq 3$.

(1) Ha F -nek nincs olyan részcsaládja, amely eleget tesz (b.t.t.)-nek, akkor bármely W_F esetén $D_F = D_{FW}$.

(2) Legyen P az $N-1$ db $S_{j-1} = \{\alpha_1, \dots, \alpha_j\}$ $j = 2, \dots, N$ részhalmazokból álló család. Bármely W_F esetén

$$(8) \quad D_F = D_{FW} = D_{FB}.$$

A tételnek számos közvetlen következménye van. Nyilvánvaló például az, hogy a fenti típusú családok esetén a választási eredmények tetszőlegesen kaotikusak lehetnek. Mivel $D_F = D_{FW}$, ezért bármely szó előáll valamely választói profilként. Tehát a korábban véltnél jóval bonyolultabb paradoxonok állnak elő; minden lehetséges! Ez azt jelenti, hogy létezik például olyan választói profil, amely az $\alpha_1 < \alpha_2 < \alpha_3$, $\alpha_3 < \alpha_4 < \alpha_5$, $\alpha_5 < \alpha_6 < \alpha_1$ választó sorrendet adja. (Ez a család nem tesz eleget (b.t.t.)-nek.) A tétel (2) állításából következik, hogy létezik olyan választói profil, amelyre $\alpha_1 > \dots > \alpha_j$, ha j páros, de $\alpha_j > \alpha_{j-1} > \dots > \alpha_1$, ha j páratlan. Az ilyen típusú példák nem a Borda módszer hibái; nincs olyan módszer a választási vektorok kiválasztására, amellyel elkerülhetők lennének ezek a negatív következmények. Bármely választás ugyanehhez a konklúzióhoz vezet. A tétel második állításának másik következménye a „kieséses” rendszerű választás standard eljárására vonatkozik. Ez az a rendszer, amelyben a végeredmény egy iterációs eljárás alapján alakul, amelynek minden lépésében kiesik az utolsó helyezett, s a megmaradt jelöltek közül újból választanak. Ez az eljárás egy monoton csökkenő (elemszámú) családot definiál, tehát természetes módon adódik, hogy a tétel második részében leírt családot vegyük figyelembe. De ekkor a tétel szerint van olyan választói profil, amelyben minden egyes S_{j-1} , $j = 2, \dots, N$ halmaz úgy van sorba rendezve, hogy α_1 minden esetben az utolsó előtti, és α_j az utolsó. (Tehát a választók többsége α_1 -t α_2 elé helyezi). Az eljárás szerint az α_1 jelölt mindig a következő fordulóba lép. Végül is ő lesz a győztes.

Ez az eredmény nem nyújtott használható módot a Borda féle számolás és más súlyokkal történő számolás közti különbségtételre. A különbségek más családok esetén merülnek fel.

2. TÉTEL (Saari, 1985 b).

- (1) Bármely F család és W_F esetén D_{FB} a D_{FW} részhalmaza.
- (2) Tegyük fel, hogy $F = \{S_1, S_2, \dots, S_T\}$ eleget tesz (b.t.t.)-nek. Ha van olyan W_F , hogy D_{FB} a D_{FW} valódi részhalmaza, akkor minden $W_F \neq B_F$ választás esetén D_{FB} valódi részhalmaza D_{FW} -nek. Ha $\sum(|S_j| - 1) > N(N - 1)/2$, akkor F rendelkezik ezzel a tulajdonsággal. Nevezetesen $N > 3$ esetén ez igaz az összes $(2^N - (N + 1))$ részhalmazt tartalmazó családra.
- (3) Egy F családhoz tartozik egy E_F Euklideszi tér; ez az a tér, amely tartalmazza az összes W_F vektort. E_F -nek van egy olyan nyílt, sűrű C_F részhalmaza, amelyre teljesül, hogy $D_F = D_{FW}$ akkor és csak akkor, ha W_F a C_F -ben van. Többek között, ha W_F választási vektor minden komponense megfelel a standard választási rendszernek, akkor W_F a C_F -ben van.

- (4) Ha $N > 3$, és az F család tartalmazza a jelöltekből képzett összes párt és az $\{\alpha_1, \dots, \alpha_N\}$ halmast, akkor C_F tartalmazza az összes vektort, kivéve B_F -et.

A tétel bizonyításában W_F -et paraméterként kezeljük. Amint azt az eredmények mutatják, egy bizonyos szingularitásfajta bukkan fel. Amint az várható, a következtetések hasonlóak — jellegüket tekintve — a szingularitáselméletben találhatóakhoz. Például, ezek a szingularitások a választási vektorok halmazára meghatároznak egy sztratifikációt. A fent értelmezett C_F halmaz határának struktúrája adja meg ezen sztratifikációt. Midőn W_F -et egy sztrátumból választjuk, a D_{FW} szótárak megegyeznek.

Ez a tétel választ ad a Borda féle eljárás első bírálatára. Ha szeretnénk a kiértékelő eljárásban szereplő súlyokat úgy megválasztani, hogy minimalizáljuk a lehetséges és látszólagos ellentmondásokat a részhalmazok egy adott családján lezajló választásban, akkor az (1) és (2) bekezdések alapján az egyetlen választás a kiértékelő eljárásra a Borda-féle számlálás. A többi eljárás által definiált szótárak több szóból állnak, tehát több eltérést engednek meg a választási sorrendben. Valójában, a (3) bekezdés alapján az általános helyzet az, hogy bármely elképzelhető, kaotikus végeredmény előfordulhat. A jelenleg használatos standard választási rendszer ezek közé tartozik, amint azt a következő állítás igazolja.

1. KÖVETKEZMÉNY. Legyen $N \geq 3$ és legyen F az összes $2^N - (N + 1)$ db részhalmazból álló család. Ezen részhalmazok mindegyikéhez válasszuk ki — tetszőlegesen — a lehetséges sorrendek valamelyikét. Ekkor létezik egy olyan választói profil (azaz a választók ki tudják tölteni úgy a szavazócédulákat), hogy ezen részhalmazok mindegyikének standard sorrendje a kiválasztott sorrend.

A kieséses, stb. rendszerekre vonatkozik az alábbi következtetés.

2. KÖVETKEZMÉNY. Legyen $N \geq 3$. Ekkor van olyan választói profil, hogy a többségi győztes a több mint 2 jelöltet tartalmazó valamennyi részhalmazban az utolsó helyre van rangsorolva. Ilyen profil esetén tehát a többségi győztes bármely kieséses rendszerű választás esetén, amelyben az utolsó helyezetteket rendre törlik, veszíteni fog. Úgysszintén létezik olyan választói profil, hogy minden három, vagy annál több jelöltet tartalmazó részhalmaz esetén a legnépszerűtlenebb lesz az első helyen rangsorolt.

A Borda-módszer második bírálatá az a tényen alapszik, hogy a módszer nem mindig rangsorolja az első helyre a többségi jelöltet. A 2. Tétel (4) bekezdéséből következik, hogy nincsen olyan súlyozott választási rendszer, amely azt tenné minden esetben. Sőt, valójában, bármely más W_F választása esetén $D_{FW} = D_F$. Következésképpen, bármely kaotikus kimenet (eredmény) lehetséges.

Az alábbiakban tekintsünk éppen egy ilyen jellegű kimenetelt.

3. KÖVETKEZMÉNY. Legyen $N \geq 3$ és s_1 és s_2 1 és N közötti, különböző természetes számok. Ekkor bármely — nem Borda típusú súlyozott választási rendszer esetén létezik olyan választói profil, hogy a többségi jelölt az s_1 -edik, a legnépszerűtlenebb jelölt az s_2 -dik helyre rangsorolódik.

A bevezető példában már láthattuk azt a szélsőséges esetet, amikor $s_1 = N$ és $s_2 = 1$. Találhatunk egyéb kritikákat is a Borda módszerre. Például úgy, hogy találunk egy olyan — a jelölteket tartalmazó részhalmazokból álló — családot, amelyre a Borda féle választási sorrend ellentmondásosnak tűnik. A 2. Tétel második részéből következik, hogy bármilyen W_F választása esetén van olyan választói profil, amely ugyanazt az eredményt adja. Ennek az az oka, hogy bármely F és W_F választás esetén D_{FB} a D_{FW} részhalmaza. Tehát a Borda módszer a legkonzisztensebb (legellentmondásmentesebb) a súlyozott választási eljárások közül.

Mivel a választási eljárásokkal probléma van, ezért természetes irányként vetődik fel, hogy csak azzal törődjünk, hogy mi történik a jelöltekkel, ha a páronkénti összevetés alapján számolják ki a köztük levő sorrendet. A kérdésnek ezt a megközelítési módját a választások elméletében fontos szerepet játszó XVIII. századbeli másik francia matematikus, Condorcet támogatta. Például ő szorgalmazta, hogy a többségi jelölt legyen győztesnek kikiáltva (CONDORCET, 1785). Egy esetleges másik megközelítésben egy forgatókönyv szerint bonyolítják le a választási gyűlést. Ez a következőképpen zajlik: Készítenek egy listát, amely tartalmazza az N jelöltet (legyen $[\alpha_1, \alpha_2, \dots, \alpha_N]$), és a lista szerint páronként összehasonlítják a jelölteket. Először az első két jelöltet, majd a többségi győztest a listán soron következő (harmadik) jelölttel vetik össze, stb. Bármely páronkénti összevetésen alapuló eljárás vezethet ellentmondásokhoz; ezt illusztrálja az alábbi két következmény. (Ezek az eredmények az összes $N(N-1)/2$ jelöltet tartalmazó családon alapulnak. Ennek a családnak nincs olyan részcsaládja, amely kielégítené (b.t.t.)-t.)

4. KÖVETKEZMÉNY. Legyen $N \geq 3$ és legyen F a jelölteket tartalmazó $N(N-1)/2$ db kételemű részhalmaz összességéből álló család. Legyen adott egy W_F által definiált választási eljárás, amely minden egyes részhalmazra többségi választási módot értelmez, ekkor $D_{FW} = D_F$. Azaz a jelöltek rangsorolásának összes lehetséges ciklusa és részciklusa előállítható a többségi választási módszer alkalmazásával.

5. KÖVETKEZMÉNY.

- (1) Ha $N \geq 3$, akkor van olyan választói profil, hogy $\alpha_j > \alpha_{j+1}$, de $\alpha_N > \alpha_1$. A többi, az előbbieken fel nem sorolt pár esetében $\alpha_j > \alpha_{j+k}$ akkor és csak akkor, ha k páratlan, pozitív egész, és $\alpha_j < \alpha_{j+k}$ akkor és csak akkor, ha k páros, pozitív egész szám.
- (2) Ezen választói profil esetén nincsen sem többségi, sem legnépszerűtlenebb jelölt.
- (3) Ezen választói profil esetén tetszőleges j -re α_j a győztes jelölt, ha az $[\alpha_{j+1}, \alpha_{j+2}, \dots, \alpha_{j-1}, \alpha_j]$ forgatókönyv szerint zajlik a választógyűlés.

Léteznek élesebb következtetések is választási eredményekről, melyeknek az érdeklődő olvasó utánanézhethet (Saari, 1985 b). Például jellemezhető a D_{FB} halmaz abból a szempontból, hogy a következő típusú eredmények közül melyek találhatók meg benne (vessük össze ezeket az eredményeket a fentiekkel).

3. TÉTEL. *A Borda-módszer az egyetlen súlyozott választási eljárás, amely sohasem rangsorolja az utolsó helyre a többségi győztest, és sohasem teszi az első helyre a legnépszerűtlenebb jelöltet.*

Lehetséges a D_{FB} és a D_{FW} halmazok elemzése tetszőleges F és W_F választás esetén. Ily módon az összes választási eredményt megtalálhatjuk.

Utolsó illusztrációként tegyük fel, hogy $N = 6$ és F a $2^6 - 7$ db legalább 2 jelöltet tartalmazó részhalmazok összességének családja. A 2. Tétel második részéből következik, hogy bármely W_F választása esetén D_{FB} valódi részhalmaza D_{FW} -nek. De tegyük fel, hogy D_{FW} csak néhány szóval tartalmaz többet, mint D_{FB} ; ekkor a fenti eredmények, amelyek a Borda-módszert támogatták, sokat veszítenek a jelentőségükből. Végül is ez azt jelenti, hogy a Borda módszer alkalmazása során kapott inkonzisztenciák száma ekkor csak kis mértékben növekedne. De a mi esetünk nem ilyen; bármely $W_F \neq B_F$ választás esetén D_{FB} lényegesen kevesebb szót tartalmaz, mint D_{FW} . Tegyük fel például, hogy F egyetlen részhalmazán sem a Borda módszerrel alakítjuk ki a sorrendet. Ekkor a messze legjobb W_F választás, amely előfordulhat, a D_{FB} minden szavának több, mint 10^{18} szavat feleltet meg a D_{FW} -ben.

Zárszó

Ez a cikk az írásos változata annak az előadásnak, amely 1985. szeptemberében hangzott el Sopronban (Magyarország), egy az IIASA-SDS által szponzorált konferencián. Ez a kutatás részben az „NSF IST 8415348” által volt támogatva. A munka egy része az IIASA-SDS-nél Ausztriában 1984. szeptemberében tett látogatásom során készült.

IRODALOM

- [1] ARROW, K., „Social Choice and Individual Values”, Monograph 12, Cowles Foundation for Research in Economics (New Have, CT, 1963).
- [2] BORDA, J.-C. DE, *Mémoire sur les élections au scrutin* (Histoire de l'Academie Royale des Sciences, Paris, 1781).
- [3] CONDORCET, MARQUIS DE, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Paris, 1785).
- [4] DEVANEY, R., *An Introduction to Chaotic Dynamical Systems* (Benjamin, Menlo Park, CA, 1986).
- [5] SAARI, D.G., „The ultimate of chaos resulting from weighted voting systems”, *Advances in Applied Mathematics* 5 (1984), 286–304.
- [6] SAARI, D.G., „Price dynamics, social choice, voting methods, probability, and chaos”, *Advances in Equilibrium Theory* (C. Aliprantis, O. Burinshaw, and N. Rothman, eds.), Lecture Notes

- in *Economics and Mathematical Systems* (Springer, Berlin, Heidelberg, New York, Tokyo, **244** (1985a)).
- [7] SAARI, D.G., „The Optimal Ranking Method is the Borda Count”, *Collaborative Paper CP-85-4* (International Institute for Applied Systems Analysis, Laxenburg, Austria, 1985b).

NORTHWESTERN EGYETEM
EWANSTON
ILLINOIS 60201
USA

Fordította: Kovács Kálmán

MATEMATIKAI PROBLÉMÁK ÉS SEJTÉSEK NUMERIKUS MEGKÖZELÍTÉSBEN*

RICHARD S. VARGA

Kent

Tartalom

Előszó

Bevezetés

1. fejezet

Az approximációelméleti Bernstein sejtés

2. fejezet

Az „ $1/9$ ” sejtés és megoldása

3. fejezet

A Riemann hipotézis elméleti és számítástechnikai szempontjai

4. fejezet

Az e^z részletösszegei zérushelyeinek asszimptotikája

5. fejezet

Valós és komplex legjobb racionális approximáció

6. fejezet

A Jensen egyenlőtlenség általánosításai alacsonyfokú koncentrációjú polinomokra

*A fordítás RICHARD S. VARGA, *Scientific Computation on Mathematical Problems and Conjectures* című könyvéből a szerző és a kiadó engedélyével készült. Ez a rész a könyv első három fejezete, a négy, öt és hatodik fejezetek fordítása a 16 (1992) 3–4-es számban lesz olvasható.

Előszó

Richard S. Varga amerikai matematikus neve nem ismeretlen számos magyar szakember, különösen a numerikus matematikával és approximációelmélettel foglalkozók számára. Bár kivándorolt szülők gyermekeként Clevelandben született, és látogatóként is csak rövid időket töltött Magyarországon, jól beszéli a magyar nyelvet, és több hazai matematikussal tart fenn szakmai-baráti kapcsolatot.

1928-ban született, Ph.D.-jét a Harvard University-n védte meg J. L. Walsh professzor irányítása alatt. 1954-60 között a Westinghouse Electric Corporation-nál dolgozott alkalmazott matematikusként (feltételezhetően innen ered a munkásságát végig kísérő, alkalmazott matematika iránti állandó érdeklődése). Egyetemi pályafutását 1960-ban a Case Institute of Technology-n kezdte; jelenlegi munkahelyén, a Kent State University-n 1970 óta dolgozik. több mint egy tucat egyetemen volt meghívott professzor (többek közt Pekingben, Valparaisóban, valamint Humboldt ösztöndíjjal Karlsruheban és Würzburgban). Oktatói és tudományos tevékenységéért számos kitüntetést kapott. Eddig 24 Ph.D. diákja volt, 9 nemzetközileg ismert szakfolyóirat szerkesztőségének tagja, és több amerikai tudományos bizottságban fejt ki tevékenységet. 15 konferencia szervezésében vett részt, 5 monográfiát írt, négyet szerkesztett, és közel 200 dolgozata jelent meg.

A jelen fordítás legújabb, a SIAM-nál megjelent, *Scientific Computation on Mathematical Problems and Conjectures* című könyvéből készült. A szerző régi vágya teljesül első magyar nyelvű közleményének megjelenésével, és meggyőződésünk, hogy a téma iránt érdeklődő olvasó sok hasznos információt nyer az anyagból.

Bevezetés

A könyvben tárgyalt előadási jegyzetekkel az a célunk, hogy bizonyos részletességgel tanulmányozzuk a tudományos számítástechnika felhasználását néhány matematikai probléma és sejtés megoldásában. Tudományos számítástechnikán elsősorban olyan számításokat értünk, melyeket *nagyszámú* szignifikáns számjeggyel végzük. E számítások a legkülönbözőbb numerikus matematikai módszerekkel kapcsolatban lépnek fel, mint például a (második) Remez algoritmus a polinomiális és racionális approximáció elméletében, a Richardson-féle számsorozat-extrapolációban (e számsorozatok konvergenciájának gyorsítására), magas fokszámú polinomok gyökeinek pontos meghatározásában, és integrálok kvadratura eljárásokkal történő numerikus approximációjában. Sietve hozzátesszük, hogy nem célunk annak a speciális témának a részletes tárgyalása, amely e nagypontosságú számításokhoz szükséges robusztus és megbízható szoftverek *megalkotásával* foglalkozik, hanem inkább annak *hangsúlyozása*, hogy a létező szoftverek mily hatalmas erőt kölcsönöznek a matematikusok fegyverzenájához, amellyel problémákat és sejtéseket támadnak meg.

Ennek szellemében az 1. fejezetben az 1913-ból származó, a polinomiális approximációra vonatkozó Bernstein sejtést tárgyaljuk. Ennek nagypontosságú számításokkal való (negatív értelmű) megoldása 1985-ből származik. A 2. fejezetben a racionális approximációra vonatkozó, 1977-ben megfogalmazott „1/9” sejtést tárgyaljuk. Ez esetben a nagypontosságú számítások erősen sugallták, hogy e sejtés sem igaz. Ennek ellenére a közelmúltban talált csodálatosan szép végső, pontos megoldás (A. A. Gonchar és E. A. Rakhmanov) inkább az elméletet, mint a numerikus matematikát gazdagítja.

A 3. fejezetben röviden áttekintjük az 1859-ből származó híres Riemann hipotézissel kapcsolatos nagypontosságú számításokat, melyek a Riemann-féle ζ -függvény kritikus sávbeli zérushelyeire vonatkoznak. Ezután az 1927-ből származó Pólya-sejtéssel foglalkozunk, amely a Riemann hipotézis egy gyengébb formája. E fejezetben (pozitív értelemben) egy analitikus megoldást adunk a Pólya-sejtésre, numerikus számításoktól vezérelve. Majd a Λ de Bruijn-Newman konstans alsó korlátainak numerikus meghatározásával zárjuk e fejezetet. Látni fogjuk, hogy e konstansra $\Lambda \leq 0$ teljesül, ha a Riemann hipotézis igaz. 1976-ból származik C. M. Newman azon sejtése, hogy $\Lambda \geq 0$.

A további fejezetekben elsősorban három, az analízisből származó elméleti problémával foglalkozunk. E fejezetekben a nagypontosságú számítások más jellegűek, ugyanis egyrészt analitikus eredményekre vonatkozó sejtések megfogalmazásához szükségesek, másrészt magasfokú polinomok gyökeinek ilyen meghatározásával érdekes grafikus képeket kapunk.

Mint látni fogjuk, e monográfiákban a hangsúlyt erősen a „kemény” analízis és a nagypontosságú számítások közti kölcsönhatásra helyeztük. Itt jegyezzük meg, hogy mindegyik fejezet független a többitől. Irodalomjegyzék minden egyes fejezet végén található.

Megköszönöm barátaim és kollégáim számos megjegyzését és javaslatát az itt

bemutatott anyaggal kapcsolatban. A Butler University-n tartott előadásokon a szerző meggondolatlanul egy-egy fánkot ajánlott fel minden egyes sajtóhibáért, amely a CBMS-NSF konferencián szétozott anyagban található. (Később sokszor gondoltam arra, hogy talán a sütőiparba kellett volna mennem!) Úgyszintén köszönöm Ms. Gail Bosticnak (Institute for Computational Mathematics, Kent State University) a kézirat kiváló \LaTeX gépelését, valamint Gretchen M. Vargának az anyag szerkesztésében végzett munkáját.

Végül mindenkifelett köszönöm Amos J. Carpenter professzornak nemcsak a Butler University-n szervezett konferencia alapos és minden részletre kiterjedő megszervezését, hanem azt a figyelmet, amellyel e könyvet olvasta; köszönöm továbbá az 1. és 4. fejezet ábráinak megrajzolását. Mint mindig, most is öröm volt vele dolgozni.

1. FEJEZET

Az approximációelméleti Bernstein sejtés

1.1. Bevezetés

Bármely valós változós, a $[-1, +1]$ intervallumon definiált $f(x)$ függvényre a folytonossági modulus tetszőleges $\delta > 0$ esetén az

$$(1.1) \quad \omega(\delta; f) := \sup_{\substack{|x_1 - x_2| \leq \delta \\ x_1, x_2 \in [-1, +1]}} |f(x_1) - f(x_2)|,$$

formulával definiáljuk, míg a $[-1, +1]$ intervallumon vett egyenletes normát az

$$(1.1') \quad \|f\|_{L_\infty[-1, +1]} := \sup \{|f(x)| : x \in [-1, +1]\}.$$

képlet adja meg. Jelölje π_n a legfeljebb n -edfokú valós polinomok halmazát ($n = 0, 1, \dots$). Jól ismert Jackson [8] (lásd Meinardus [9, 56. oldal], Rivlin [11, 22. oldal]) alábbi eredménye:

1. TÉTEL (Jackson [8]). Ha $f(x)$ egy folytonos, valós értékű, a $[-1, +1]$ intervallumon definiált függvény, akkor

$$(1.2) \quad E_n(f) \leq 6\omega\left(\frac{1}{n}; f\right) \quad (n = 1, 2, \dots),$$

ahol

$$(1.3) \quad E_n(f) := \inf \{\|f - g\|_{L_\infty[-1, +1]} : g \in \pi_n\}.$$

Ismeretes (lásd [9, 16. oldal]), hogy bármely, a $[-1, +1]$ intervallumon definiált folytonos, valós értékű $f(x)$ függvényhez létezik egy *egyértelmű* $\hat{p}_n(x) = \hat{p}_n(x; f)$ a π_n -ben úgy, hogy

$$(1.3') \quad E_n(f) = \|f - \hat{p}_n\|_{L_\infty[-1, +1]} \quad (n = 0, 1, \dots),$$

és ezt a $\hat{p}_n(x)$ -et az $f(x)$ π_n -ből vett *legjobb egyenletes approximációjának* nevezzük a $[-1, +1]$ -en. Továbbá világos (1.3)-ból, hogy $\{E_n(f)\}_{n=0}^\infty$ nemnegatív számok egy nem növekvő sorozata, amely Weierstrass approximációelméleti tétele (lásd [11, 11. oldal]) értelmében zérushoz tart, vagyis

$$(1.4) \quad \lim_{n \rightarrow \infty} E_n(f) = 0.$$

Nevezetesen, az $|x|$ függvényre a $[-1, +1]$ intervallumon könnyen látható, hogy

$$\omega(\delta; |x|) = \delta \quad (0 < \delta \leq 1),$$

úgyhogy az 1. Tétel (1.2) formulája szerint

$$(1.4') \quad E_n(|x|) \leq \frac{6}{n} \quad (n = 1, 2, \dots),$$

ami (1.4)-nek egy, az $|x|$ -re vonatkozó pontosabb formája. Mivel $|x|$ egy páros folytonos függvény a $[-1, +1]$ -en, ezért a π_n -ből vett legjobb egyenletes approximációja a $[-1, +1]$ -en is ilyen, bármely $n \geq 0$ esetén (lásd Rivlin [11, 43. oldal, Exercise 1.1]). E megfigyelést összevetve a legjobb egyenletes polinomiális approximáció Csebisev alternációs jellemzésével, következik (lásd [11, 26. oldal]), hogy

$$(1.5) \quad E_{2n}(|x|) = E_{2n+1}(|x|) \quad (n = 0, 1, \dots).$$

Ezért számunkra elegendő azt vizsgálni, hogy az $\{E_{2n}(|x|)\}_{n=1}^{\infty}$ sorozat hogyan tart (csökkenőleg) zérushoz. Az (1.4')-ből nyilvánvalóan következik, hogy

$$(1.6) \quad 2nE_{2n}(|x|) \leq 6 \quad (n = 1, 2, \dots).$$

Az (1.6)-beli felső korlátot javítandó, fejtsük $|x|$ -et Csebisev sorba a $[-1, +1]$ -en, vagyis

$$(1.7) \quad |x| = \frac{4}{\pi} \left\{ \frac{1}{2} + \sum_{m=1}^{\infty} \frac{(-1)^{m+1} T_{2m}(x)}{(2m-1)(2m+1)} \right\},$$

ahol $T_n(x)$ jelöli az n -ik (elsőfajú) Csebisev polinomot. Az (1.7)-beli összeg első n tagját véve, az (1.7)-beli maradékösszeg kielégíti a

$$\frac{4}{\pi} \left| \sum_{m=n+1}^{\infty} \frac{(-1)^{m+1} T_{2m}(x)}{(2m-1)(2m+1)} \right| \leq \frac{4}{\pi} \sum_{m=n+1}^{\infty} \frac{1}{(2m-1)(2m+1)},$$

egyenlőtlenséget, mivel $|T_n(x)| \leq 1$ bármely $[-1, +1]$ -beli x esetén. Másrészt az

$$\frac{1}{(2m-1)(2m+1)} = \frac{1}{2} \left(\frac{1}{2m-1} - \frac{1}{2m+1} \right)$$

összefüggés miatt ez a felső korlát teleszkópikusan a

$$\begin{aligned} \frac{4}{\pi} \sum_{m=n+1}^{\infty} \frac{1}{(2m-1)(2m+1)} &= \frac{2}{\pi} \left\{ \left(\frac{1}{2n+1} - \frac{1}{2n+3} \right) + \right. \\ &\quad \left. + \left(\frac{1}{2n+3} - \frac{1}{2n+5} \right) + \dots \right\} \\ &= \frac{2}{\pi(2n+1)} \end{aligned}$$

mennyiségre egyszerűsödik. Így

$$(1.8) \quad \left| |x| - \frac{4}{\pi} \left\{ \frac{1}{2} + \sum_{m=1}^n \frac{(-1)^{m+1} T_{2m}(x)}{(2m-1)(2m+1)} \right\} \right| \leq \frac{2}{\pi(2n+1)},$$

bármely $[-1, +1]$ -beli x és minden $n \geq 1$ esetén. Mivel $|x|$ (1.8)-beli approximációja egy $2n$ -edfokú polinom, (1.8) maga után vonja, hogy

$$(1.9) \quad 2nE_{2n}(|x|) \leq \frac{4n}{\pi(2n+1)} < \frac{2}{\pi} = 0,63661 \dots \quad (n = 1, 2, \dots),$$

ami jobb, mint (1.6).

S. Bernstein [2] 1913-ban jelentősen megjavította az (1.9)-beli $0,63661 \dots$ felső korlátot. Míg (1.9) közvetlenül a

$$(1.10) \quad \overline{\lim}_{n \rightarrow \infty} 2nE_{2n}(|x|) \leq \frac{2}{\pi} = 0,63661 \dots$$

becslést adja, Bernstein, egy hosszú és nehéz bizonyítással az alábbi, sokkal mélyebb eredményt kapta:

2. TÉTEL (Bernstein [2]). *Létezik egy β pozitív konstans (β Bernstein nevének kezdőbetűje miatt) úgy, hogy*

$$(1.11) \quad \lim_{n \rightarrow \infty} 2nE_{2n}(|x|) = \beta,$$

ahol β kielégíti az alábbi egyenlőtlenséget:

$$(1.12) \quad 0,278 < \beta < 0,286.$$

Ezen túlmenően Bernstein [2, 56. oldal], mint kuriózumot jegyezte meg, hogy az

$$(1.13) \quad \frac{1}{2\sqrt{\pi}} = 0,28209\,47917 \dots$$

konstans ugyancsak az (1.12)-beli korlátok közé esik, és nagyon jó közelítéssel e β -ra vonatkozó korlátok *aritmetikai közepe*, nevezetesen $0,282$. E megfigyelés az évek során úgy vált ismertté, mint a

$$(1.14) \quad \text{Bernstein sejtés (1913): } \beta \stackrel{?}{=} \frac{1}{2\sqrt{\pi}} = 0,28209\,47917 \dots$$

A Bernstein munkájának megjelenése óta eltelt több mint 70 év alatt e sejtés megoldatlan maradt, bár több szerző numerikus úton próbálkozott a megoldással (lásd Bell és Shah [1], Bojanic és Elkins [3], valamint Salvati [12]). E hosszú időn

át tartó sikertelenség oka valószínűleg az, hogy (i) az $E_{2n}(|x|)$ számok pontos meghatározása nagy n esetén numerikus szempontból *nemtriviális*, és (ii) a $2nE_{2n}(|x|)$ konvergenciája β -hoz (amelyet (1.11) garantál), egészen *lassú*.

A közelmúltban, 1985-ben Varga és Carpenter [13] bebizonyította, hogy a Bernstein sejtés *nem igaz*; ez következik az alábbi, β -ra vonatkozó élesebb korlátokból [13] (ezeket a következő paragrafusokban tárgyaljuk):

$$(1.15) \quad 0,28016\,85460\dots = \ell_{20} \leq \beta \leq 2\mu_{100} = 0,28017\,33791\dots$$

Mivel az (1.15)-beli, β -ra vonatkozó felső korlát *kisebb*, mint $1/(2\sqrt{\pi}) = 0,28209\,47917\dots$, az (1.14) alatti Bernstein sejtés *nem igaz*! Alább az 1.2 paragrafusban $E_{2n}(|x|)$ nagypontosságú meghatározásával foglalkozunk, míg az 1.3 és 1.4 részekben röviden tárgyaljuk a [13]-beli számításokat, amelyek Bernsteinnek a β konstansra vonatkozó becslési módszerein alapulnak. 1.5-ben a nagypontosságú $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számok Richardson extrapolációjával foglalkozunk, ami a β -ra az alábbi, 50 számjegyre pontos becslést adja:

$$(1.16) \quad \begin{array}{r} \beta \doteq 0,28016\,94990\,23869\,13303\,64364 \\ 91230\,67200\,00424\,82139\,81236, \end{array}$$

s végül 1.6-ban egy új sejtést mondunk ki a $2nE_{2n}(|x|)$ (6.3)-beli aszimptotikus sorbafejtéséről, megemlítve néhány kapcsolódó nyitott problémát.

1.2 Az $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számok nagypontosságú kiszámítása

Legyen a π_{2n} -beli $\hat{p}_{2n}(x)$ az $|x|$ egyértelmű legjobb egyenletes approximációja a π_{2n} -ből a $[-1, +1]$ -en; vagyis (lásd (1.3'))

$$(2.1) \quad \||x| - \hat{p}_{2n}(x)\|_{L_\infty[-1, +1]} = E_{2n}(|x|) \quad (n = 1, 2, \dots).$$

Az 1.1-beli megfontolások alapján, mivel $|x|$ páros a $[-1, +1]$ -en, ezért legjobb egyenletes approximációja a π_{2n} -ből a $[-1, +1]$ intervallumon is ilyen, amiből következik, hogy

$$(2.2) \quad \hat{p}_{2n}(x) = \sum_{j=0}^n a_j(n)x^{2j} \quad (n = 1, 2, \dots).$$

Bevezetve az $x^2 = t$, $t \in [0, 1]$ új változót, approximációs problémánk az

$$(2.3) \quad E_{2n}(|x|) = E_n(\sqrt{t}; [0, 1]) := \inf \left\{ \|\sqrt{t} - h_n(t)\|_{L_\infty[0, 1]} : h_n \in \pi_n \right\}$$

alakot ölti. Mármost, ha

$$(2.4) \quad E_n(\sqrt{t}; [0, 1]) = \|\sqrt{t} - \hat{h}_n(t)\|_{L_\infty[0, 1]} \quad (\hat{h}_n \in \pi_n),$$

akkor nyilván (lásd (2.2))

$$(2.5) \quad \hat{p}_{2n}(x) = \hat{h}_n(x^2) \quad (n = 1, 2, \dots).$$

Így az $E_{2n}(|x|)$ és $\hat{p}_{2n}(x)$ értékek meghatározása ekvivalens az $E_n(\sqrt{t}; [0, 1])$ és $\hat{h}_n(t)$ kiszámításával.

A (2.3) minimalizálási problémát a (második) Remez algoritmus alábbi, meglehetősen standard implementációjával oldhatjuk meg (lásd Meinardus [9, 105. oldal]):

1. lépés. Legyen $S := \{t_j\}_{j=0}^{n+1}$ $n+2$ különböző pont halmaza a $[0, 1]$ intervallumon úgy, hogy

$$(2.6) \quad 0 \leq t_0 < t_1 < \dots < t_{n+1} \leq 1.$$

2. lépés. Keressük meg azt az egyértelmű $h_n(t)$ polinomot és a λ konstanst (ami egy lineáris probléma) úgy, hogy

$$(2.7) \quad h_n(t_j) + (-1)^j \lambda = \sqrt{t_j} \quad (j = 0, 1, \dots, n+1).$$

Így $h_n(t)$ a \sqrt{t} legjobb egyenletes approximációja a π_n -ből e diszkrét S halmazon, alternálva $|\lambda|$ hibával az S szukcesszív t_j pontjaiban. Így a (2.3) jelöléssel összhangban azt írhatjuk, hogy

$$(2.8) \quad \|\sqrt{t} - h_n(t)\|_{L_\infty(S)} = E_n(\sqrt{t}; S) = |\lambda|.$$

Mivel S a $[0, 1]$ egy részhalmaza, nyilván

$$(2.9) \quad \|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]} - |\lambda| \geq 0.$$

3. lépés. Egy előre megadott (kis) $\varepsilon > 0$ -val, ha $\|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]} - |\lambda| \leq \varepsilon$, akkor az iteráció végetért. Különben pedig keressünk egy új S' halmazt a $\sqrt{t} - h_n(t)$ függvény $[0, 1]$ -beli lokális extrémumaiból, váltakozó előjelekkel, a 2. lépés alkalmazásával, s ismételjük a 2. és 3. lépéseket, stb. Numerikus szempontból hasznos tudni, hogy ezen Remez algoritmus ismételt alkalmazásával nyert $|\lambda|$ sorozat monoton *növekedő*. Kiindulva a speciális $S^{(0)} := \{t_j^{(0)}\}_{j=0}^{n+1}$ alternáló halmazból, ahol

$$(2.10) \quad t_j^{(0)} := \frac{1}{2} \left\{ 1 + \cos \left[\frac{(n+1-j)\pi}{n+1} \right] \right\} \quad (j = 0, 1, \dots, n+1)$$

pontok a $[0, 1]$ intervallumbeli $T_{n+1}(2t-1)$ Csebisev polinom $n+2$ extrémális pontjai, és felhasználva Brent [4] többszörös pontosságú algoritmusát a Kent State University, Department of Mathematical Sciences VAX 11/780 számítógépén, a Remez algoritmus iteráltjai akkor álltak be, amikor $\|\sqrt{t} - h_n(t)\|_{L_\infty[0,1]}$ és $|\lambda|$ 100 értékes jegyre megegyeztek (lásd (2.9)). E (második) Remez algoritmus ismert kvadrátikus konvergenciája miatt (lásd [9, 113. oldal]), legfeljebb *kilenc* iterációra volt

n	$2nE_{2n}(x)$	n	$2nE_{2n}(x)$
1	0,25000000000000000000	27	0,28010923652220618525
2	0,27048359711113710107	28	0,28011346088995028384
3	0,27557437240117538604	29	0,28011725624949961792
4	0,27751782467505269646	30	0,28012067877266282833
5	0,27845118553550860152	31	0,28012377573166088450
6	0,27896791746495870636	32	0,28012658713873191844
7	0,27928294495851802460	33	0,28012914704390451720
8	0,27948883759450744771	34	0,28013148457001261069
9	0,27963065741012820125	35	0,28013362474403004676
10	0,27973243377197382968	36	0,28013558916927111713
11	0,27980791728874387383	37	0,28013739657233669662
12	0,27896543212379327279	38	0,28013906325078289591
13	0,27991025431555769036	39	0,28014060344158248218
14	0,27994585848578213247	40	0,28014202962599794087
15	0,27997460668640749231	41	0,28014335278310408169
16	0,27999815195631672827	42	0,28014458260161108707
17	0,28001767713329725379	43	0,28014572765764550097
18	0,28003404741499350964	44	0,28014679556460041624
19	0,28004790728590585156	45	0,28014779309995913546
20	0,28005974476042315265	46	0,28014872631304874446
21	0,28006993483180943067	47	0,28014960061693143684
22	0,28007876947528753423	48	0,28015042086704695023
23	0,28008647875707557049	49	0,28015119142874492326
24	0,28009324593880850547	50	0,28015191623546527355
25	0,28009921845238283558	51	0,28015259883901781632
26	0,28010451598655670489	52	0,28015324245316384249

1.1. Táblázat $\{2nE_{2n}(|x|)\}_{n=1}^{52}$

szükség bármelyik vizsgált esetben. Egy kissé túlbiztosítva a számításokat, és tekintetbe véve lehetséges kerekítési hibákat is, meggyőződésünk, hogy az általunk meghatározott $\{E_{2n}(|x|)\}_{n=1}^{52}$ számok legalább 95 értékes jegyre pontosak.

A rövidség kedvéért a $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ szorzatokat 20 értékes jegyre kerekítve adjuk meg az 1.1. Táblázatban, hogy érzékeltesük e sorozat lassú konvergenciáját. (A $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ kinyomtatott értékei 100 értékes jegyre igény szerint rendelkezésre állnak.)

Úgy tűnik, hogy az 1.1. Táblázatbeli $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ szorzatok négy értékes jegyre konvergálnak. Az 1.6.-beli aszimptotikus becslések alapján, a $|2nE_{2n}(|x|) - \beta| < 10^{-10}$ egyenlőtlenséghez $n \geq 20\,968$ szükséges. Ehhez viszont az kellene, hogy megkeressük \sqrt{t} legjobb egyenletes, legalább 10 484-edfokú approximációját a $[0, 1]$ intervallumon, ami valóban egy félelmetesen nagy feladat lenne!

1.3. A Bernstein-féle β konstans felső korlátainak kiszámítása

Abból a célból, hogy felső és alsó korlátokat kapjon a β konstansra, Bernstein [2] bevezette a speciális

$$(3.1) \quad F(t) := \sum_{k=0}^{\infty} \frac{t}{(t+2k+1)^2 - \frac{1}{4}} \quad (t \geq 0)$$

függvényt. A

$$(3.2) \quad \Psi(z) := \frac{d}{dz}(\log \Gamma(z)) = \frac{\Gamma'(z)}{\Gamma(z)}$$

függvény segítségével (lásd Whittaker and Watson [14, 240. oldal]) a (3.1)-beli $F(t)$ a következő alakban írható:

$$(3.3) \quad F(t) = \frac{t}{2} \left\{ \Psi\left(\frac{t}{2} + \frac{3}{4}\right) - \Psi\left(\frac{t}{2} + \frac{1}{4}\right) \right\} \quad (t \geq 0).$$

Az $F(t)$ egy másik reprezentációja (lásd [2])

$$(3.4) \quad F(t) = \frac{t}{2t+1} F\left(1, 1; t + \frac{3}{2}; \frac{1}{2}\right) = t \int_0^1 \frac{z^{t-1/2} dz}{z+1} = \frac{1}{2} \int_0^{\infty} \frac{e^{-u} du}{\cosh(u/2t)},$$

ahol $F(a, b; c; z)$ a klasszikus hipergeometrikus függvény (lásd Henrici [7, 56. oldal]). A (3.4)-beli utolsó integrál mutatja, hogy $F(t)$ szigorúan növekedő a $[0, +\infty)$ intervallumon, $F(0) = 0$ és $F(+\infty) = \frac{1}{2}$.

Ezen a ponton még egyáltalán nem világos, hogy a (3.1)-beli $F(t)$ függvény miért játszana szerepet a β Bernstein konstans meghatározásában. E függvény szerepének megvilágítására Bernstein a következő polinomiális interpolációs problémát tekintette. Minden egyes rögzített n egész számra tekintsük a $2n+1$ különböző

$$(3.5) \quad x_0 := x_0(2n) = 0; \quad x_k := x_k(2n) := \cos \left[\frac{(k-1/2)\pi}{2n} \right] \quad (k = 1, 2, \dots, 2n)$$

pontot a $[-1, +1]$ intervallumban. Mivel az $\{x_k\}_{k=1}^{2n}$ pontok a $T_{2n}(x)$ Csebisev polinom zérushelyei, ezért

$$(3.6) \quad \omega(x) := \prod_{j=0}^{2n} (x - x_j) = x T_{2n}(x) / 2^{2n-1}.$$

Ha $R_{2n}(x)$ jelöli azt az egyértelmű, π_{2n} -beli polinomot, amely interpolálja az $|x|$ függvényt a (3.5) alatti $2n+1$ pontban, akkor $R_{2n}(x)$ az x változónak egy páros polinomja, amely egy kis átalakítás után kielégíti az

$$(3.7) \quad |x| - R_{2n}(x) = \frac{T_{2n}(x)}{n} H_{2n}(x) \quad (x \in [0, 1])$$

relációt, ahol

$$(3.7') \quad H_{2n}(x) := - \sum_{k=0}^{n-1} \frac{(-1)^k \sin[(k + \frac{1}{2}) \frac{\pi}{2n}]}{x + \cos[(k + \frac{1}{2}) \frac{\pi}{2n}]} \quad (x \in [0, 1]).$$

Egy hosszú bizonyítással Bernstein [2] megmutatta, hogy

$$(3.8) \quad |x| - R_{2n}(x) = \frac{T_{2n}(x)}{n} \left[F\left(\frac{2nx}{\pi}\right) + \eta_n(x) \right] \quad (x \in [0, 1]),$$

ahol

$$(3.9) \quad |\eta_n(x)| \leq \frac{4 + \pi^2}{2n^{2/5}} \quad (x \in [0, 1]; \quad n = 1, 2, \dots).$$

Mármost definíció szerint

$$E_{2n}(|x|) \leq \| |x| - R_{2n}(x) \|_{L_\infty[-1, +1]} = \| |x| - R_{2n}(x) \|_{L_\infty[0, 1]},$$

ahol az utolsó egyenlőséget az indokolja, hogy $|x|$ és $R_{2n}(x)$ egyaránt páros függvények. Mivel $|T_{2n}(x)| \leq 1$ a $[0, 1]$ -en, és mivel $F(t)$ egy növekvő függvény a $[0, +\infty)$ intervallumon úgy, hogy $F(+\infty) = 1/2$, ezért

$$(3.10) \quad \overline{\lim}_{n \rightarrow \infty} 2n E_{2n}(|x|) \leq 2F(+\infty) = 1.$$

Jegyezzük meg, hogy (3.10), miközben javítja (1.6)-ot, nem olyan jó, mint az (1.10) alatti eredmény. Továbbá (3.8) mutatja, hogy az $|x| - R_{2n}(x)$ hibák *messze* nem ekvioszcillálnak a $[-1, +1]$ intervallumban nagy n -re, mivel $F(t)$ szigorúan növekvő volta maga után vonja, hogy a legnagyobb ilyen hibák csak az $x = \pm 1$ környezetében fordulhatnak elő, nem pedig $x = 0$ környezetében. Abból a célból, hogy az $x = 0$ környezetében előidézzük az ekvioszcillációt, Bernstein [2] a következőt javasolta. Először is jegyezzük meg, hogy bármely páros n esetén

$$(3.11) \quad T_{2n}(x) = \cos(2n \arccos x) = \cos(2n \arcsin x).$$

A továbbiakban rögzített nemnegatív m egész és bármely $n > m$ esetén, jelölje $\left\{ \xi_k(2n) := \sin \left[\frac{(2k-1)\pi}{4n} \right] \right\}_{k=1}^m$ a $T_{2n}(x)$ polinom m legkisebb gyökét a $[0, 1]$ intervallumban. A $T_{2n}(x) / (x^2 - \xi_k^2(2n))$ nyilván egy páros polinom a π_{2n} -ből minden egyes k -ra, hacsak $1 \leq k \leq m$, úgyhogy a

$$(3.12) \quad Q_{2n}(x) := R_{2n}(x) + \frac{T_{2n}(x)}{n} \left\{ a_0 + \left(\frac{\pi}{2n} \right)^2 \sum_{k=1}^m \frac{a_k}{x^2 - \xi_k^2(2n)} \right\}$$

polinom szintén páros és π_{2n} -ből való minden $n > m$ esetén. Továbbá (3.8) és (3.12) maga után vonja, hogy $x = \pi b/(2n)$ esetén

$$(3.13) \quad |x| - Q_{2n}(x) = \frac{T_{2n}(x)}{n} \left\{ F(b) - \left(a_0 + \sum_{k=1}^m \frac{a_k}{b^2 - \left[\frac{2n}{\pi} \sin\left(\frac{(2k-1)\pi}{4n}\right) \right]^2} \right) + \eta_n(x) \right\},$$

hacsak $n > m$. Legyen

$$(3.14) \quad \mu_m := \inf_{a_0, a_1, \dots, a_m \text{ real}} \left\{ \left\| \cos(\pi b) \left[F(b) - \left(a_0 + \sum_{k=1}^m \frac{a_k}{b^2 - \left(\frac{2k-1}{2}\right)^2} \right) \right] \right\|_{L_\infty[0, +\infty)} \right\},$$

bármely nemnegatív m egész esetén. Mivel (3.11) szerint bármely rögzített $b \geq 0$ esetén

$$T_{2n}(x) = \cos \left[2n \arcsin \left(\frac{\pi b}{2n} \right) \right] \rightarrow \cos(\pi b) \quad (n \rightarrow \infty),$$

a (3.10)-hez analóg módon belátható, hogy

$$(3.15) \quad \overline{\lim}_{n \rightarrow \infty} 2n E_{2n}(|x|) \leq 2\mu_m \quad (m = 0, 1, \dots).$$

A $\{\mu_m\}_{m=0}^\infty$ pozitív konstansok sorozata a (3.14)-ben nyilván nemnövekvő:

$$(3.16) \quad \mu_0 \geq \mu_1 \geq \mu_2 \geq \dots,$$

és így konvergens. Mármost Bernstein [2, 55. oldal] bebizonyította, hogy az (1.11)-beli β konstans, valamint a $\{\mu_m\}_{m=0}^\infty$ sorozat között fennáll a

$$(3.17) \quad \beta = 2 \lim_{m \rightarrow \infty} \mu_m$$

reláció. Így (3.16) alapján, a (3.14) approximációs probléma bármely μ_m konstansa a következő felső korlátot szolgáltatja β -ra:

$$(3.18) \quad \beta \leq 2\mu_m \quad (m = 0, 1, \dots).$$

Érdekes módon 1913-ban Bernstein [2] numerikusan megbecsülte a (3.14) megoldását $m = 3$ esetén, és azt találta, hogy $\mu_3 < 0,143$, úgyhogy

$$2\mu_3 < 0,286,$$

ami β (1.12) alatt említett felső korlátja. A $2\mu_m$ pontosabb becslései az alábbi 1.2 táblázatban találhatók. Viszont a (3.14) alatti minimizációs probléma, vagyis a

μ_m konstans meghatározása nem más, mint az $F(b)$ speciális súlyozott racionális approximációja a $[0, +\infty)$ intervallumon, ahol a $\cos(\pi b)$ súly (lásd (3.14)) természetesen nem állandó előjelű ezen az intervallumon. Mint arra [13]-ban rámutattunk, kiderült, hogy a (3.14) alatti approximációs probléma megoldása egy érdekes oszcillációval jellemezhető, amely megengedi a (második) Remez algoritmus módosított formájának használatát. (Meg kell jegyeznünk, hogy Bernstein [2] 1913-beli munkája megelőzi a [10] alatti Remez algoritmus 1934-es közlését!) A (3.14) alatti minimalizálási problémát az alábbi módosított (második) Remez algoritmus segítségével oldották meg:

1. lépés. Legyen $\tilde{S} := \{t_j\}_{j=0}^{m+1}$ ($m \geq 1$) egy $m+2$ különböző pontból álló halmaz a $[0, +\infty)$ intervallumon, amely kielégíti az alábbi egyenlőtlenségeket:

$$(3.19) \quad 0 = t_0 < t_1 < \dots < t_m \leq m - \frac{1}{2} < t_{m+1} =: +\infty.$$

2. lépés. Keressük meg azt a $m+2$ egyértelmű $\{a_i\}_{i=0}^m$ és $|\lambda|$ konstanst, amely kielégíti a

$$(3.20) \quad \begin{cases} \cos(\pi t_j) \left\{ a_0 + \sum_{k=1}^m \frac{a_k}{t_j^2 - [(2k-1)/2]^2} \right\} + (-1)^{j+1} \lambda = \\ \quad = \cos(\pi t_j) F(t_j) \quad (j = 0, \dots, m), \\ a_0 + \lambda = \frac{1}{2} = F(+\infty) \end{cases}$$

feltételeket. E (3.20) lineáris probléma megoldása az

$$(3.21) \quad R_m(t) := \cos(\pi t) \left[F(t) - \left(a_0 + \sum_{k=1}^m \frac{a_k}{t^2 - [(2k-1)/2]^2} \right) \right]$$

függvény ekvioszcillációját eredményezi az \tilde{S} halmaz $\{t_j\}_{j=0}^m$ részhalmazán $|\lambda|$ alternáló hibával a szukcesszív $\{t_j\}_{j=0}^m$ pontokban, és ezenkívül arra készíti $R_m(t)$ -et, hogy $+|\lambda|$ és $-|\lambda|$ között oszcilláljon, midőn $t \rightarrow \infty$ (lásd az 1.1. ábrát). Így (2.8)-hoz hasonlóan,

$$(3.22) \quad \|R_m(t)\|_{L_\infty(\tilde{S})} = |\lambda|,$$

és, mivel \tilde{S} a $[0, +\infty)$ egy részhalmaza, ezért (lásd (2.9))

$$(3.23) \quad \|R_m(t)\|_{L_\infty[0, +\infty)} - |\lambda| \geq 0.$$

E módosított Remez algoritmus következő lépésében megjelenő feltételek értelmezése céljából megjegyezzük, hogy

$$(3.24) \quad \begin{aligned} \frac{d}{dt} \left\{ F(t) - \left(a_0 + \sum_{k=1}^m \frac{a_k}{t^2 - [(2k-1)/2]^2} \right) \right\} = \\ = F'(t) + 2t \sum_{k=1}^m \frac{a_k}{[t^2 - [(2k-1)/2]^2]^2}. \end{aligned}$$

m	2μ	m	$2\mu_m$
0	0,5000000000000000	10	0,280568148084662
1	0,309816648277486	20	0,280267918128026
2	0,289644642836759	30	0,280213001347551
3	0,284585623264382	40	0,280193895181171
4	0,282681644408752	50	0,280185082723738
5	0,281779992624272	60	0,280180306766681
6	0,281286520869723	70	0,280177431742434
7	0,280988433465837	80	0,280175568033390
8	0,280795058278109	90	0,280174291500582
9	0,280662672087176	100	0,280173379101718

 1.2. Táblázat $\{2\mu_m\}$

Így, ha (3.20) megoldása olyan, hogy az a_k ($0 \leq k \leq m$) és λ értékek mind pozitívak, akkor $F(t)$ szigorúan növekedő volta miatt a (3.24) alapján

$$G_m(t) := F(t) - \left(a_0 + \sum_{k=1}^m \frac{a_k}{t^2 - [(2k-1)/2]^2} \right)$$

szigorúan növekszik $-\infty$ -től $+\lambda$ -ig az $(m - \frac{1}{2}, +\infty)$ intervallumon. Legyen $\tau_m = \tau_m(\tilde{S})$ a t azon egyértelmű, $(m - \frac{1}{2}, +\infty)$ intervallumbeli értéke, melyre $G_m(\tau_m) = -\lambda$. Ekkor, mivel $|R_m(t)| = |\cos(\pi t) \cdot G_m(t)| \leq |G_m(t)| \leq \lambda$ minden $t \geq \tau_m$ esetén, továbbá $G_m(t) \rightarrow \frac{1}{2} - a_0 = \lambda$ midőn $t \rightarrow +\infty$, következik, hogy

$$\|R_m(t)\|_{L_\infty[\tau_m, +\infty)} = |\lambda|.$$

Másrészt, mivel $\{t_j\}_{j=0}^m$ egy részhalmaza a $[0, m - \frac{1}{2}]$ intervallumnak, (3.22) alapján $\|R_m(t)\|_{L_\infty[0, m-1/2]} \geq |\lambda|$. Innen, mivel $[0, m - \frac{1}{2}]$ részhalmaza a $[0, \tau_m]$ intervallumnak,

$$(3.25) \quad \|R_m(t)\|_{L_\infty[0, \tau_m(\tilde{S})]} = \|R_m(t)\|_{L_\infty[0, +\infty)},$$

és numerikusan eldönthető, hogy a

$$(3.25') \quad \|R_m(t)\|_{L_\infty[0, m-1/2]} = \|R_m(t)\|_{L_\infty[0, +\infty)}$$

erősebb állítás igaz-e vagy sem. (A β felső korlátainak speciális eseteit tárgyaló eddigi irodalomban, úgymint $m = 1$ Bojanic és Elkins [3], valamint $m = 3$ Bernstein [2] esetében, a hozzárendelt \tilde{S} halmazok olyanok voltak, hogy a 2. lépésbeli a_k ($0 \leq k \leq m$) és λ értékek pozitívak voltak, és (3.25') teljesült.) Ezzel elérkeztünk a 3. lépéshez.

9. lépés. Egy előre megadott (kis) $\varepsilon > 0$ mellett, ha a 2. lépésbeli $\{a_k\}_{k=0}^m$ és λ mind pozitívak, ha (3.25') teljesül, és ha

$$\|R_m(t)\|_{L_\infty[0, m-1/2]} - \lambda \leq \varepsilon,$$

akkor az iteráció véget ér. Máskülönben keresünk egy új, az $R_m(t)$ függvény m (váltakozó előjelű) lokális extremumaiból álló \tilde{S}' halmazt a $(0, m - \frac{1}{2})$ intervallumon a $t_0 := 0$ és $t_{m+1} := \infty$ pontokhoz, és megismételjük a 2. és 3. lépéseket, stb. Kiindulva az $\tilde{S}^{(0)} := \{t_j^{(0)}\}_{j=0}^{m+1}$ halmazból, ahol

$$(3.26) \quad t_0^{(0)} := 0; \quad t_j^{(0)} := \frac{2j-1}{2} \quad (j = 1, 2, \dots, m); \quad t_{m+1}^{(0)} := +\infty,$$

a megfelelő (3.20) lineáris egyenletrendszer megoldható az $\{a_k^{(0)}\}_{k=0}^m$ és $\lambda^{(0)}$ ismeretleneire, a Gauss-féle eliminációs módszerrel, részleges főelemkiválasztással. Mindegyik, az 1.2. Táblázatban tekintett esetben az \tilde{S} halmaz (3.26) alatti alternációs pontokból álló kezdőértékei olyannyira jók voltak, hogy a fenti módosított Remez algoritmus 2. és 3. lépései mindig előállították a pozitív $\{a_k\}_{k=0}^m$ és λ számokat, valamint a (3.19) és (3.25')-et kielégítő alternációs halmazokat. Ezenkívül, mint az várható volt, ezen algoritmus konvergenciája *kvadrátikus*, és e módosított Remez algoritmus legfeljebb 10 lépésére volt szükség a tekintett esetekben a konvergenciához.

Az 1.1. ábrán a (3.21)-beli $R_5(t)$ függvény gráfját ábráztuk, amelyhez a (3.14)-beli legjobb approximációs μ_5 konstans tartozik. Ezen az ábrán hat alternációs pont látható (pontokkal megjelölve) a $[0, 9/2]$ intervallumban, valamint a $(9/2, +\infty)$ intervallumbeli oszcillációk, amelyek abszolút értékben a μ_5 értékhez növekednek, midőn $t \rightarrow \infty$.

Sajnos, a fenti számításokat nem végeztük el ugyanazzal a nagy pontossággal (100 értékes jegyre), amivel az $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ 1.2-beli számításait. Ennek egyik oka, hogy a (3.14)-re alkalmazott módosított Remez algoritmus szükségszerűen igényli a (3.1)-beli $F(t)$ függvény ismételt kiszámítását. Ehhez az $F(t)$ (3.3) alatti, a $\Psi(x)$ függvény segítségével megadott reprezentációját használtuk, és $\Psi(x)$ -et egyrészt a

$$(3.27) \quad \Psi(x) \doteq (x - x_0)r_{8,8}(x) \quad \left(\frac{1}{2} \leq x \leq 3\right),$$

függvénnyel approximáltuk (lásd Cody, Strecok és Thacher [7]), ahol $x_0 = 1.46163 \dots$ a $\Psi(x)$ egyetlen pozitív zérushelye, amelyet 40 értékes jegyre ismerünk, és ahol (lásd [7]) $r_{8,8}(x)$ két konkrét, x -ben 8-adfokú polinom hányadosa; másrészt használtuk a

$$(3.28) \quad \Psi(x) \doteq \ln x - \frac{1}{2x} + r_{6,6}\left(\frac{1}{x^2}\right) \quad (3 \leq x < \infty),$$

approximációt, ahol (lásd [7]) $r_{6,6}(u)$ két konkrét 6-odfokú polinom hányadosa. A $0 < x \leq \frac{1}{2}$ intervallumban az ismert

$$(3.29) \quad \Psi(x) = \Psi(x+1) - \frac{1}{x},$$

rekurzív összefüggést, valamint a (3.27) approximációt használtuk.

Mármost a $\Psi(t)$ fenti approximációja kb. 20 értékes jegyre jó (lásd [7]), úgy-hogy becslésünk szerint a $\{2\mu_m\}_{m=0}^{100}$ -re vonatkozó számításaink legalább 15 értékes jegyre pontosak. Helykimélés céljából a $\{2\mu_m\}_{m=0}^{100}$ számoknak egy 15 értékes jegyre kerekített részhalmazát adtuk meg (1.2. táblázat).

Az 1.2. Táblázatból világos, hogy (1.13) és (3.16)–(3.17) alapján

$$(3.30) \quad \frac{1}{2\sqrt{\pi}} = 0.28209\,47917\ldots > 2\mu_5 > \beta,$$

vagyis a Bernstein sejtés *nem igaz*

1.4. A Bernstein-féle β konstans alsó korlátainak kiszámítása

A Bernstein-féle β konstans $\{2\mu_m\}_{m=0}^{100}$ felső korlátainak 1.3. pontbeli számításainál láttuk (lásd (1.11)), hogy ez *elegendő* volt az (1.14) Bernstein sejtés cáfolatához. Így e vonatkozásban nyilván *szükségtelen* alsó korlátokat megadni β -ra. Mindazonáltal a teljesség kedvéért megadunk ilyen alsó korlátokat, Bernstein egy másik ötletes módszerét alkalmazva. E számítások bizonyultak a *legkényelmesebbnek*.

Bernsteinnek [2] a β alsó korlátaira vonatkozó módszerét ismerttetendő, legyen

$$(4.1) \quad \phi_m(x) := \prod_{j=1}^{m-1} (x^2 - j^2) \quad (m \geq 1),$$

és

$$(4.2) \quad \psi_m(x) = \psi_m(x; \lambda_1, \lambda_2, \dots, \lambda_m) := \prod_{j=1}^m (x^2 - \lambda_j^2) \quad (m \geq 1).$$

(Itt azt a konvenciót használtuk, hogy $\prod_{j=\beta}^{\alpha} := 1$ ha $\alpha < \beta$.) A (4.2)-beli λ_j paraméterekről feltesszük, hogy kielégítik a

$$(4.3) \quad j-1 < \lambda_j < j \quad (j \geq 1).$$

feltételeket. Így $m \geq 1$ esetén legyen

$$(4.4) \quad B_m(\lambda_1, \lambda_2, \dots, \lambda_m) := \frac{\sum_{i=1}^m \frac{\phi_m(\lambda_i)}{\psi'_m(\lambda_i)} \left[1 - \left(\frac{2\lambda_i}{\lambda_i + \frac{1}{2}} \right) F\left(\lambda_i + \frac{1}{2}\right) \right]}{\sum_{i=1}^m \frac{\phi_m(\lambda_i)}{\psi'_m(\lambda_i)} \left[\frac{2}{\pi\lambda_i} + \tan\left(\frac{\pi}{2}[\lambda_i - i + 1]\right) \right]},$$

ahol az $F(t)$ függvény a (3.1) alatt van definiálva. Jegyezzük meg, hogy a (4.1) és (4.2) alapján azt írhatjuk, hogy

$$(4.5) \quad \frac{\phi_m(\lambda_i)}{\psi'_m(\lambda_i)} = \frac{\prod_{j=1}^{i-1} (\lambda_i^2 - j^2) \cdot \prod_{j=i}^{m-1} (j^2 - \lambda_i^2)}{2\lambda_i \prod_{j=1}^{i-1} (\lambda_i^2 - \lambda_j^2) \cdot \prod_{j=i+1}^m (\lambda_j^2 - \lambda_i^2)} \quad (1 \leq i \leq m).$$

A (4.3) feltételekből látjuk, hogy a fenti hányadosok mind pozitívak. Így, mivel $F(t)$ szigorúan monoton növekedő a $[0, +\infty)$ intervallumon, $F(0) = 0$, és $F(+\infty) = \frac{1}{2}$ (lásd 1.3. pont), arra következtetünk, hogy a (4.4)-beli összegek minden egyes tagja pozitív, ahonnan $B_m(\lambda_1, \lambda_2, \dots, \lambda_m) > 0$.

Az (1.11) alatt definiált β -val Bernstein [2] megmutatta, hogy $B_m(\lambda_1, \lambda_2, \dots, \lambda_m)$ egy alsó korlát a β -ra, vagyis

$$(4.6) \quad \beta \geq B_m(\lambda_1, \lambda_2, \dots, \lambda_m)$$

bármely pozitív egész m és a (4.3)-at kielégítő bármely $\{\lambda_j\}_{j=1}^m$ esetén. A legjobb ilyen alsó korlát β -ra bármely $m \geq 1$ esetén nyilván

$$(4.7) \quad l_m := \max \{ B_m(\lambda_1, \lambda_2, \dots, \lambda_m) : \{\lambda_j\}_{j=1}^m \text{ kielégíti (4.3)-at} \},$$

úgyhogy

$$(4.8) \quad \beta \geq l_m > 0 \quad (m \geq 1).$$

Mármost tekintsünk bármely, a (4.3)-at kielégítő $\{\lambda_j\}_{j=1}^{m+1}$ paramétereket. Rögzítve a $\{\lambda_j\}_{j=1}^m$ értékeket, és λ_{m+1} -et az m -ig csökkentve, láthatjuk a (4.4) és (4.5)-ből, hogy

$$(4.9) \quad \lim_{\lambda_{m+1} \rightarrow m} B_{m+1}(\lambda_1, \lambda_2, \dots, \lambda_m, \lambda_{m+1}) = B_m(\lambda_1, \lambda_2, \dots, \lambda_m).$$

Következésképpen, a (4.7) alapján

$$(4.10) \quad l_{m+1} \geq l_m \quad (m \geq 1),$$

vagyis (4.8) miatt $\{l_m\}_{m=1}^\infty$ pozitív számoknak egy korlátos, nemcsökkenő sorozata. Bernstein [2] ugyancsak megmutatta, hogy e sorozat határértéke β :

$$(4.11) \quad \beta = \lim_{m \rightarrow \infty} l_m.$$

Bernstein [2] numerikusan megbecsülte az l_1 és l_2 értékeket, és azt találta, hogy

$$(4.12) \quad l_1 > 0,27 \quad \text{és} \quad l_2 > 0,278.$$

m	l_m	m	l_m
1	0,271982359030477	11	0,280163464187524
2	0,278930922849406	12	0,280164893327009
3	0,279811000437231	13	0,280165905238063
4	0,280024333928903	14	0,280166641527680
5	0,280097791315214	15	0,280167189892928
6	0,280129183079687	16	0,280167606600825
7	0,280144691009336	17	0,280167928871653
8	0,280153187711753	18	0,280168181990114
9	0,280158217699044	19	0,280168383539180
10	0,280161379471687	20	0,280168546002042

1.3. Táblázat $\{\ell_m\}_{m=1}^{20}$

Az l_2 ezen utóbbi becslése, mint a β alsó korlátja mutatkozik (1.12)-ben. (Az l_1 és l_2 pontosabb becslése található az 1.3. Táblázatban.)

Ezekután ismertetjük az l_m alsó korlátaira vonatkozó számításainkat. A (4.4)-ből világos, hogy a $\lambda_1, \lambda_2, \dots, \lambda_m$ paraméterek nem lineárisan lépnek fel a $B_m(\lambda_1, \lambda_2, \dots, \lambda_m)$ definíciójában. Számítástechnikailag egy meglehetősen standard optimizációs (maximalizáló) rutint használtunk, deriváltak nélkül, a (4.3) korlátozásoknak alávetett $\{\lambda_i\}_{i=1}^m$ valós paraméterek meghatározására, amelyek maximalizálták a $B_m(\lambda_1, \lambda_2, \dots, \lambda_m)$ értéket, és így megadták l_m (4.7) alatti értékét. Úgyisntén, mivel az $F(t)$ függvény explicite szerepel a $B_m(\lambda_1, \lambda_2, \dots, \lambda_m)$ (4.4) alatti definíciójában, felhasználtuk a $\Psi(t)$ függvény (3.27)–(3.29) alatti approximációját, valamint az $F(t)$ (3.3) alatti reprezentációját a $\Psi(t)$ -vel kifejezve. Csakúgy, mint a β (3.17) alatti $2\mu_m$ felső korlátainak kiszámításánál, a β konstans l_m alsó korlátainak kiszámításánál sem követtük a $2nE_{2n}(|x|)$ 1.2. pontbeli kiszámításánál alkalmazott nagyfokú pontosságot (95 értékes jegy). Az 1.3. pont numerikus eredményeihez hasonlóan, mivel a $\Psi(t)$ -re vonatkozó approximációnk kb. 20 értékes jegyre pontos (lásd [7]), úgy gondoljuk, hogy az $\{l_m\}_{m=1}^{20}$ értékekre vonatkozó optimizációs számítások legalább 15 értékes jegyre pontosak. A 15 értékes jegyre kerekített $\{l_m\}_{m=1}^{20}$ számokat az 1.3. Táblázatban adtuk meg.

Ha az 1.2. Táblázat felső korlátait összehasonlítjuk az 1.3. Táblázat alsó korlátával, akkor láthatjuk, hogy a (4.8)-beli l_m alsó korlát *lényegesen* pontosabb becslése az (1.16) alatti β értéknek, mint a (3.17) alatti $2\mu_m$ felső korlát, *bármely* $1 \leq m \leq 20$ esetén. Valóban, a β -t approximáló l_{20} hibája durván $9,53 \cdot 10^{-7}$, míg ugyanez $2\mu_{100}$ esetén csak $3,88 \cdot 10^{-6}$. Mindazonáltal a pontosság ezen növekedésnek nagyjából el-lensúlyozza a *megegyezkedett* gépidő, amely az l_m számoknak az általunk alkalmazott optimizációs rutinnal való megtalálásához szükséges. A β -t approximáló l_m alsó korlátok e nagyobb pontossága magyarázza azt, hogy az 1.3. Táblázatbeli m értékek *miért* nem vesznek fel oly magas értékeket, mint az 1.2. Táblázatbeli $2\mu_m$ esetén.

1.5. A $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számok Richardson-féle extrapolációja

Az 1.1. Táblázatban megjelenő $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számok jelzik, hogy ezek konvergenciája a β Bernstein-konstanshoz (lásd (1.11)) meglehetősen lassú. A lassan konvergáló sorozatok konvergencia-sebességének javítására kidolgozott módszerek tipikus képviselője az ún. *Richardson-féle extrapolációs módszer* (lásd Brezinski [5, 7. oldal]), amely a következőképpen írható le. Ha $\{S_n\}_{n=1}^N$, a valós számok egy adott (véges) sorozata, ahol $N > 2$, akkor legyen $T_0^{(n)} := S_n$ ($1 \leq n \leq N$), és tekintsük a $\{T_0^{(n)}\}_{n=1}^N$ értékeket mint a Richardson-féle extrapolációs táblázat N számból álló nulladik oszlopát. A Richardson-féle extrapolációs táblázat $N - 1$ elemből álló első oszlopát a

$$(5.1) \quad T_1^{(n)} := \frac{x_n T_0^{(n+1)} - x_{n+1} T_0^{(n)}}{x_n - x_{n+1}} \quad (1 \leq n \leq N - 1),$$

formulákkal definiáljuk, és induktíve, az $N - k - 1$ elemből álló $(k + 1)$ -edik oszlop legyen

$$(5.2) \quad T_{k+1}^{(n)} := \frac{x_n T_k^{(n+1)} - x_{n+k+1} T_k^{(n)}}{x_n - x_{n+k+1}} \quad (1 \leq n \leq N - k - 1),$$

minden $k = 0, 1, \dots, N - 2$ esetén, ahol $\{x_n\}_{n=1}^N$ adott konstansok. Ily módon egy $N(N + 1)/2$ elemből álló, háromszög alakú táblázatot konstruáltunk. Esetünkben a $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számokból kiindulva, egy 1378 elemből álló háromszög alakú táblázatot szerkeszthetünk. Ami az (5.1)–(5.2) alatti $\{x_n\}_{n=1}^{52}$ számok választását illeti, az előzetes számítások azt mutatták, hogy $2nE_{2n}(|x|) \doteq \beta + K/n^2 + \text{alacsonyabbrendű tagok}$, úgyhogy az $x_n := 1/n^2$ értéket választottuk. Megjegyezzük, hogy az (5.1) és (5.2) alatt definiált törtek számlálóiban és nevezőiben fellépő kivonások okozta lehetséges pontosságvesztések azt *sugallták*, hogy a $2nE_{2n}(|x|)$ kiszámítását nagyon pontosan (95 értékes jegyre) kell elvégezni.

A $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számok Richardson-féle extrapolációja *váratlanul szép eredményeket hozott*. Ahelyett, hogy bemutatnánk az 1 378 elem teljes extrapolációs táblázatát (megadva mindegyik elemet mondjuk 95 értékes jegyre), elég megemlíteni, hogy e táblázat utolsó 20 oszlopában szereplő 210 elem 3 kivételével megegyezett az (1.16)-beli β alábbi, 45 számjegyre megadott értékével:

$$(5.3) \quad \begin{aligned} \beta &\doteq 0,28016\ 94990\ 23869\ 13303\ 64364 \\ &\quad 91230\ 67200\ 00424\ 82139\ 81236. \end{aligned}$$

Ezenkívül e Richardson-féle extrapolációs táblázat 20 utolsó oszlopának 210 eleméből 182 elem olyan, hogy az első ötven számjegy *mind* megegyezett a β (5.3) alatti approximációjával!

1.6. Néhány nyitott probléma

A $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számokra $x_n := 1/n^2$ értékkel alkalmazott Richardson-féle extrapoláció sikere határozottan sejteti, hogy $2nE_{2n}(|x|)$ rendelkezik egy

$$(6.1) \quad 2nE_{2n}(|x|) \stackrel{?}{\approx} \beta - \frac{K_1}{n^2} + \frac{K_2}{n^4} - \frac{K_3}{n^6} + \dots \quad (n \rightarrow \infty)$$

alakú aszimptotikus kifejtéssel (lásd Henrici [7, 355. oldal]), ahol a K_j konstansok függetlenek az n -től. Feltételezve, hogy (6.1) igaz, következik, hogy

$$(6.2) \quad n^2(2nE_{2n}(|x|) - \beta) \approx -K_1 + \frac{K_2}{n^2} - \frac{K_3}{n^4} + \dots \quad (n \rightarrow \infty).$$

Így az 1.1. Táblázatbeli $2nE_{2n}(|x|)$ értékek ismert nagypontosságú approximációjával, valamint a β -nak egy a $\{2nE_{2n}(|x|)\}_{n=1}^{52}$ számokra vonatkozó Richardson-féle extrapolációs táblázata utolsó eleméből vett becslésével ismét alkalmazhatjuk a Richardson-féle extrapolációt az $\{n^2(2nE_{2n}(|x|) - \beta)\}_{n=1}^{52}$ számokra (az $x_n = 1/n^2$ értékkel), nyervén egy extrapolált becslést a (6.2)-beli K_1 értékre. Ez a „szorosabbra fűzéses” eljárás folytatható a Richardson-féle extrapoláció segítségével oly módon, hogy szukcesszív becsléseket nyerjünk a (6.1)-beli K_j értékekre. Mint az gyanítható, a K_j konstansok e szukcesszív meghatározása a numerikus pontosság fokozatos csökkenéséhez vezet.

j	K_j
1	0,04396 75288 8
2	0,02640 71687 7
3	0,03125 34264 6
4	0,05889 00165 7
5	0,16010 69971
6	0,59543 53151
7	2,92591 5470
8	18,49414 033
9	146,94301 23
10	1438,03271 7

1.4. Táblázat $\{K_j\}_{j=1}^{10}$

Az 1.4. Táblázatban a $\{K_j\}_{j=1}^{10}$ értékekre vonatkozó, 10 értékes jegyre kerekített becsléseket tüntettük fel. Mint azt a táblázat jelzi, a K_j konstansok egész gyorsan növekednek. Mivel ezek a konstansok mind *pozitívnak* bizonyultak, kimondjuk az alábbi új sejtést.

SEJTÉS (Varga és Carpenter [13]). A $2nE_{2n}(|x|)$ rendelkezik egy

$$(6.3) \quad 2nE_{2n}(|x|) \stackrel{?}{\approx} \beta - \frac{K_1}{n^2} + \frac{K_2}{n^4} - \frac{K_3}{n^6} + \dots \quad (n \rightarrow \infty)$$

formájú aszimptotikus kifejtéssel, ahol a K_j konstansok (amelyek n -től függetlenek) mind pozitívak.

Mármost, mivel a Bernstein-féle β konstans kapcsolatban van az $F(t)$ függvény bizonyos racionális approximációival (lásd (3.16)), ahol $F(t)$ kifejezhető (lásd (3.4)) egy klasszikus hipergeometrikus függvény segítségével, nem indokolatlan azt gondolni, hogy β , valamint a (6.3)-beli K_j konstansok megadhatók klasszikus hipergeometrikus függvények és/vagy ismert matematikai konstansok zárt kifejezéseiként.

Végül, Bernstein [2] 1913-ban írott dolgozatának végén azt állítja, hogy az $E_{2n}(|x|; [-1, +1])$ hiba viselkedésének analízise kiterjeszthető az $E_{2n}(|x|^{2\alpha}; [-1, +1])$ viselkedésének vizsgálatára, bármely $\alpha > 0$ mellett, vagyis hogy létezik egy pozitív $\beta(\alpha)$ konstans úgy, hogy

$$(6.4) \quad \beta(\alpha) = \lim_{n \rightarrow \infty} (2n)^{2\alpha} E_{2n}(|x|^{2\alpha}; [-1, +1]) = \lim_{n \rightarrow \infty} (2n)^{2\alpha} E_n(x^\alpha; [0, 1]) \quad (\alpha > 0),$$

ahol az utolsó egyenlőség — csakúgy, mint (2.3) esetén — következik az $|x|^{2\alpha} [-1, +1]$ intervallumon való párosságából. Így e fejezet előző részei a fenti jelölésben kizárólag az (5.3) által adott $\beta(1/2)$ becslésére irányultak.

Érdekes módon, 25 évvel [2] dolgozatának 1913-as megjelenése után Bernstein visszatért e probléma vizsgálatára, mégpedig az 1938-ban írott [20] és [21] dolgozatában a $\beta(\alpha)$ függvény néhány figyelemreméltó tulajdonságával kapcsolatban. Számtalan érdekes, de nehéz módszert alkalmazva, Bernstein [20] megtalálta a

$$(6.5) \quad \beta(\alpha) < \frac{\Gamma(2\alpha)|\sin(\pi\alpha)|}{\pi} \quad (\alpha > 0)$$

felső, valamint a

$$(6.6) \quad \frac{\Gamma(2\alpha)|\sin(\pi\alpha)|}{\pi} \left(1 - \frac{1}{2\alpha - 1}\right) < \beta(\alpha) \quad \left(\alpha > \frac{1}{2}\right)$$

alsó korlátot, amelyekből következik az alábbi pontos aszimptotikus viselkedés, mielőtt $\alpha \rightarrow \infty$:

$$(6.7) \quad \lim_{\alpha \rightarrow \infty} \frac{\beta(\alpha)}{\Gamma(2\alpha)|\sin(\pi\alpha)|/\pi} = 1.$$

A [20]-ban ugyancsak bebizonyította, hogy

$$(6.8) \quad \lim_{\alpha \downarrow 0} \beta(\alpha) = \frac{1}{2} =: \beta(0).$$

A (6.5)-(6.8) alatti eredményeket úgy is felfoghatjuk, mint Bernstein egy kísérletét arra vonatkozólag, hogy a $\beta(\alpha)$ függvényt zárt formában előállítsa $\alpha \geq 0$ esetén. Ám, eltekintve a (6.8)-beli $\beta(0)$, valamint az (5.3)-beli $\beta(1/2)$ értékektől, semmilyen más $\beta(\alpha)$ érték sem ismeretes az irodalomban. Mindazonáltal újabban Carpenter

α	$\beta(\alpha)$
0, 000	5, 00000 00000 00000 00000 00000 00000 00000 00000 00000E - 01
0, 125	3, 92106 06865 24306 18102 87889 06500 29516 74073E - 01
0, 250	3, 48648 23272 56100 43273 50066 60904 27053 37181E - 01
0, 375	3, 15241 27414 61107 18764 66738 56654 82499 32994E - 01
0, 500	2, 80169 49902 38691 33036 43649 12306 72000 04248E - 01
0, 625	2, 36444 76483 36463 84095 46777 48284 40668 05347E - 01
0, 750	1, 78360 33169 26983 67018 81533 55271 40747 72851E - 01
0, 875	1, 00876 79735 91345 44168 60888 21616 98102 46483E - 01
1, 000	0, 00000 00000 00000 00000 00000 00000 00000 00000E + 00
1, 250	2, 74965 97507 96998 83110 43751 61314 34800 90980E - 01
1, 500	5, 91106 95862 73252 18719 17623 47676 77064 56234E - 01
1, 750	7, 00436 70509 81847 47876 41483 52647 24759 44023E - 01
2, 000	0, 00000 00000 00000 00000 00000 00000 00000 00000E + 00
2, 250	2, 48295 48477 65220 75909 82753 63988 73708 98876E + 00
2, 500	7, 28031 92389 13027 78383 71440 91741 94820 71406E + 00
2, 750	1, 12733 90258 80507 55856 52198 44242 73647 99596E + 01
3, 000	0, 00000 00000 00000 00000 00000 00000 00000 00000E + 00

1.5. Táblázat:

Becslések $\beta(\alpha)$ értékeire 40 tizedesjegyre

és Varga [22], az előző részekhez hasonló nagy pontosságú számítások és extrapolációk alkalmazásával meghatározták az alábbi, legalább 40 tizedesjegyre pontos becsléseket $\beta(\alpha)$ -ra, az 1.5. Táblázatban megadott α értékekre.

Az e probléma megoldására irányuló összes elméleti és számítástechnikai munka ellenére, $\beta(\alpha)$, $\alpha > 0$ zárt formában való előállítása egy nyitott probléma marad!

1.7. Az $|x|$ racionális approximációja a $[-1, 1]$ intervallumon

Mivel e fejezet előző részeiben a Bernstein sejtéssel foglalkoztunk, vagyis az $|x|$ $[-1, +1]$ intervallumon való egyenletes *polinomiális* approximációjával, természetes e fejezet végén az $|x|$ *racionális* approximációját vizsgálni a $[-1, +1]$ intervallumon. Bármely nemnegatív egész n esetén jelentse $\pi_{n,n}$ az összes $r_{n,n}(x) = p(x)/q(x)$ racionális függvény halmazát, ahol $p \in \pi_n$ és $q \in \pi_n$. (Itt feltesszük, hogy p és q nem rendelkeznek közös tényezővel, q nem tűnik el a $[-1, +1]$ intervallumon, és q úgy van normalizálva, hogy $q(0) = 1$.) Ekkor az (1.3) analógiájára bármely valós értékű, a $[-1, +1]$ intervallumon definiált $f(x)$ függvényhez hozzárendelhetjük az

$$(7.1) \quad E_{n,n}(f) := \inf \{ \|f - r_{n,n}\|_{L_\infty[-1, +1]} : r_{n,n} \in \pi_{n,n} \}.$$

mennyiséget.

Érdekes, hogy míg Bernstein [2] 1913-ban teljes mélységében vizsgálta az $|x|$ legjobb egyenletes *polinomiális* approximációjának viselkedését a $[-1, +1]$ intervallumon, addig csak ötven évvel később, 1964-ben D. J. Newman [12] mutatott rá

arra, hogy az $|x|$ *racionális* approximációja a $[-1, +1]$ intervallumon mily alapvetően *különböző*, amennyiben Newman konstruktív módon megmutatta, hogy

$$(7.2) \quad \frac{1}{2e^{9\sqrt{n}}} \leq E_{n,n}(|x|) \leq \frac{3}{e^{\sqrt{n}}} \quad (n = 4, 5, \dots).$$

A (7.2) alatti Newman-féle egyenlőtlenségek sok kutató érdeklődését felkeltették, és Bernstein korábbi, az $E_n(|x|)$ ($n \rightarrow \infty$) aszimptotikus viselkedésére vonatkozó vizsgálatainak szellemében e kutatások jelentős része a megfelelő analóg problémára koncentrált, vagyis az $E_{n,n}(|x|)$ éles aszimptotikus viselkedésére, midőn $n \rightarrow \infty$.

Az $E_{n,n}(f)$ aszimptotikus viselkedésének általános elméletében fontos előrelépést jelentett Gonchar [8] és mások munkája. Konkrétan az $E_{n,n}(|x|)$ mennyiséggel kapcsolatban a jelenlegi legjobb eredmény Bulanov [6] nevéhez fűződik, aki bebizonyította, hogy

$$(7.3) \quad E_{n,n}(|x|) \geq e^{-\pi\sqrt{n+1}} \quad (n = 0, 1, \dots).$$

Ezenkívül Vjacheslavov [18] megmutatta, hogy léteznek M_1 és M_2 pozitív konstansok úgy, hogy

$$(7.4) \quad M_1 \leq e^{\pi\sqrt{n}} E_{n,n}(|x|) \leq M_2 \quad (n = 1, 2, \dots).$$

A (7.3) és (7.4) nyilván maga után vonja, hogy

$$(7.5) \quad e^{\pi(1-\sqrt{2})} = 0,27218 \dots \leq e^{\pi\sqrt{n}} E_{n,n}(|x|) \leq M_2 \quad (n = 1, 2, \dots),$$

és ha

$$(7.6) \quad \underline{M} := \lim_{n \rightarrow \infty} e^{\pi\sqrt{n}} E_{n,n}(|x|) \text{ és } \overline{M} := \overline{\lim}_{n \rightarrow \infty} e^{\pi\sqrt{n}} E_{n,n}(|x|),$$

akkor

$$(7.7) \quad 1 \leq \underline{M} \leq \overline{M}.$$

A (7.4) nyilván megadja a π -t mint a \sqrt{n} aszimptotikusan éles szorzóját az $E_{n,n}(|x|)$ aszimptotikus viselkedésében, midőn $n \rightarrow \infty$. Ezekután csak a (7.7)-beli legjobb aszimptotikus \underline{M} és \overline{M} konstansok meghatározása marad hátra.

Annak érdekében, hogy e problémát alaposabban megvizsgáljuk, most ismertetjük Varga, Ruttan és Carpenter [17] legújabb, nagypontosságú számításait az $\{E_{n,n}(|x|)\}_{n=1}^{40}$ számokkal kapcsolatban. Csakúgy, mint az 1.2. pontbeli polinomiális esetben, bármely nemnegatív n egész esetén az $|x|$ legjobb egyenletes approximációja a $[-1, +1]$ intervallumon a $\pi_{n,n}$ halmazból, mondjuk $\hat{r}_{n,n}(x)$, egyértelmű (lásd [11, 158. oldal]), s így

$$(7.8) \quad E_{n,n}(|x|) = \| |x| - \hat{r}_{n,n}(x) \|_{L_\infty[-1, +1]} \quad (n = 1, 2, \dots).$$

Továbbá, mivel $|x|$ páros a $[-1, +1]$ intervallumon, $\hat{r}_{n,n}(x)$ is ilyen, s ezért megmutatható (lásd (1.5)), hogy

$$(7.9) \quad E_{2n,2n}(|x|) = E_{2n+1,2n+1}(|x|) \quad (n = 1, 2, \dots).$$

Mint eddig is, célunkhoz elegendő az $\{E_{2n,2n}(|x|)\}_{n=0}^{\infty}$ sorozat zérushoz tartásának módját vizsgálni.

Mármost ha $\hat{h}_{n,n}(t) \in \pi_{n,n}$ a \sqrt{t} legjobb egyenletes approximációja a $[0, 1]$ intervallumon a $\pi_{n,n}$ halmazból bármely $n = 1, 2, \dots$ esetén, vagyis ha

$$(7.10) \quad E_{n,n}(\sqrt{t}; [0, 1]) := \inf_{r_{n,n} \in \pi_{n,n}} \|\sqrt{t} - r_{n,n}(t)\|_{L_{\infty}[0,1]} = \|\sqrt{t} - \hat{h}_{n,n}(t)\|_{L_{\infty}[0,1]},$$

akkor könnyen megmutatható (lásd (2.4)), hogy

$$(7.11) \quad E_{2n,2n}(|x|) = E_{n,n}(\sqrt{t}; [0, 1]) \quad (n = 1, 2, \dots),$$

ahol

$$(7.12) \quad \hat{r}_{2n,2n}(x) = \hat{h}_{n,n}(x^2) \quad (n = 1, 2, \dots).$$

A (7.12)-ből az $\{E_{2n,2n}(|x|)\}_{n=1}^{40}$ becsléseit közvetlenül az $\{E_{n,n}(\sqrt{t}; [0, 1])\}_{n=1}^{40}$ nagypontosságú kiszámításából kaptuk. E számításoknál — csakúgy, mint a 1.2. pontban — a (második) Remez algoritmust használtuk, s itt Brent [4] nagypontosságú programját alkalmaztuk 250 értékes számjegyre, tartalék számjegyet tartva, és megengedve kis kerekítési hibákat. Úgy gondoljuk, hogy a kiszámított $\{E_{n,n}(\sqrt{t}; [0, 1])\}_{n=1}^{40}$ mennyiségek 200 értékes jegyre pontosak. Az $\{E_{2n,2n}(|x|)\}_{n=21}^{40}$ és $\{e^{\pi\sqrt{2n}} E_{2n,2n}(|x|)\}_{n=21}^{40}$ számokat az 1.6. Táblázatban adtuk meg, 25 jegyre kerekítve.

Mint az 1.5. pontban, itt is alkalmaztunk néhány különböző extrapolációs eljárást (Richardson-féle extrapoláció, Aitken-féle Δ^2 extrapoláció, stb.; lásd Brezinski [5]) az $\{e^{\pi\sqrt{2n}} E_{2n,2n}(|x|)\}_{n=1}^{21}$ számokkal kapcsolatban. A legjobb eredményeket a Richardson-féle extrapolációval értük el, az $x_n = 1/\sqrt{n}$ választással. Az $\{e^{\pi\sqrt{2n}} E_{2n,2n}(|x|)\}_{n=21}^{40}$ számok 9. és 10. Richardson extrapolációjának eredményeit az 1.7. Táblázat foglalja össze. Az előbbieket szigorúan csökkenő, az utóbbiak szigorúan növekedő sorozatot alkotnak.

E táblázat extrapolációira alapozva, Varga, Ruttan és Carpenter [17] kimondták az alábbi, numerikusan meglehetősen nyilvánvaló sejtést:

$$(7.13) \quad \lim_{n \rightarrow \infty} e^{\pi\sqrt{2n}} E_{2n,2n}(|x|) \stackrel{?}{=} 8.$$

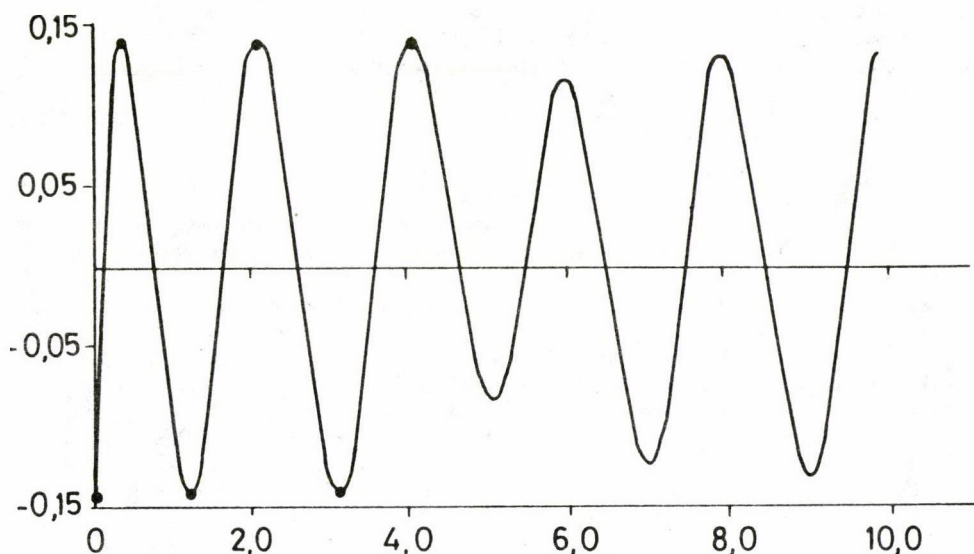
Érdekességként megemlíjtjük, hogy Herbert Stahl professzor megkérte a [17] szerzőitől az \sqrt{x} függvénynek a $[0, 1]$ intervallumon vett legjobb egyenletes racionális approximációja hibájának, vagyis az $E_{n,n}(\sqrt{x}; [0, 1])$ mennyiségeknek a szélsőértékeit. Ezek a numerikus eredmények nyilván nagyon hasznosak voltak számára, hiszen [23]-ban képes volt *elméletileg* bebizonyítani, hogy a (7.13) sejtés *helyes*! Végül, az 1.6. pontbeli tárgyalással teljes analógiában és a (7.13) szellemében természetesen vetődik fel az $E_{n,n}(x^{\alpha}; [0, 1])$ viselkedésének kérdése $n \rightarrow \infty$ esetén, bármely $\alpha > 0$ mellett. Ez megtalálható Varga és Carpenter [24] közlés alatt lévő munkájában.

n	$E_{2n,2n}(x ; [-1, +1])$	$e^{\pi\sqrt{2n}} E_{2n,2n}(x ; [-1, +1])$
21	9,6011226128422364808987184e - 9	6,6756165126491228856564179
22	5,9708233987055580552986137e - 9	6,7032142882249977256424257
23	3,7523813816413163690864502e - 9	6,7291099634760209110520998
24	2,3814996907217830892279694e - 9	6,7534733658511869861964983
25	1,5254732895109793748147207e - 9	6,7764513791852569033345348
26	9,8567633494963529958137413e - 10	6,7981717950311136695770741
27	6,4213580507266246923653248e - 10	6,8187464002912796750796788
28	4,2158848429927145758285061e - 10	6,8382734742229698180371436
29	2,7883241651339275411060214e - 10	6,8568398240938623267702634
30	1,8570720011628217953125707e - 10	6,8745224571336711172475540
31	1,2450783250744235910902360e - 10	6,8913899632991017639054615
32	8,4005997557762786343216049e - 11	6,9075036662673253080419613
33	5,7022115757288620263774447e - 11	6,9229185872920030400076656
34	3,8929505815993459443909823e - 11	6,9376842569099166681845857
35	2,6724435566456537363975894e - 11	6,9518454021392401752909853
36	1,8442995092525441602503777e - 11	6,9654425311662094614637204
37	1,2792448409247089881993010e - 11	6,9785124331456697053440800
38	8,9163582949186860871201939e - 12	6,9910886073298323319862475
39	6,2438281549962812624730424e - 12	7,0032016330585887701672461
40	4.3920484091817861898391037e - 12	7,0148794900233669056665337

1.6. Táblázat

$A \{\tau_n\}_{n=21}^{40}$ 9. Richardson extrapolációja	$A \{\tau_n\}_{n=21}^{40}$ 10. Richardson extrapolációja
8,0000000004818513852150904	7,9999999993370575957653022
8,0000000002792857242205205	7,9999999996174919169009855
8,0000000001662223537658992	7,9999999997766671415448461
8,0000000001018861846283786	7,9999999998673924596859198
8,0000000000644065954058002	7,9999999999194597844657179
8,0000000000419621984410583	7,9999999999496419688750299
8,0000000000280990775511207	7,9999999999673808389086599
8,0000000000192489204346099	7,9999999999779992400786189
8,0000000000134077625530325	7,9999999999845068292101649
8,0000000000094285808538428	7,999999999986129550248035
8,0000000000066398157884231	

1.7. Táblázat

1.1. Ábra $R_5(t)$

IRODALOM

- [1] BELL, R. A. and SHAH, S. M., „Oscillating polynomials and approximations to $|x|$ ”, *Publ. of the Ramanujan Inst.* **1** (1969), 167–177.
- [2] BERNSTEIN, S., „Sur la meilleure approximation de $|x|$ par des polynômes de degré donnés”, *Acta Math.* **37** (1913), 1–57.
- [3] BOJANIC, R. and ELKINS, J. M., „Bernstein’s constant and best approximation on $[0, \infty)$ ”, *Publ. de l’Inst. Mat., Nouvelle série* **18** (32) (1975), 19–30.
- [4] BRENT, RICHARD P., „A FORTRAN multiple-precision arithmetic package”, *Assoc. Comput. Mach. Trans. Math. Software* **4** (1978), 57–70.
- [5] BREZINSKI, C., *Algorithms d’Accélération de la Convergence* (Éditions Technip, Paris, 1978).
- [6] BULANOV, A. P., „Asymptotics for least deviation of $|x|$ from rational functions”, *Mat. Sbornik* **76** (118) (1968), 288–303, (oroszul); Angol fordítás: *Math. USSR Sbornik* **5** (1968), 275–290.
- [7] CODY, W. J., STRECOK, A. J. and THACHER, H. C. JR., „Chebyshev approximations for the psi function”, *Math. Comp.* **27** (1973), 123–127.
- [8] GONCHAR, A. A., „Estimates of the growth of rational functions and some of their applications”, *Mat. Sbornik* **72**(114) (1967), 489–503, (oroszul); Angol fordítás: *Math. USSR Sbornik* **1** (1967), 445–456.
- [9] HENRICI, P., *Applied and Computational Complex Analysis*, vol. 2 (John Wiley & Sons, New York, 1974).
- [10] JACKSON, D., „Über die Genauigkeit der Annäherung stetiger Funktionen durch rationale

- Funktionen gegebenen Grades und trigonometrische Summen gegebener Ordnung", Dissertation University of Göttingen.
- [11] MEINARDUS, G., *Approximation of Functions: Theory and Numerical Methods* (Springer-Verlag Inc., New York, 1967).
 - [12] NEWMAN, D. J., "Rational approximation to $|x|$ ", *Michigan Math. J.* **11** (1964), 11–14.
 - [13] REMEZ, E. YA., "Sur le calcul effectif des polynômes d'approximation de Tchebichef", *C.R. Acad. Sci. Paris* **199** (1934), 337–340.
 - [14] RIVLIN, T. J., *An Introduction to the Approximation of Functions* (Blaisdell Publishing Co., Waltham, Mass., 1969).
 - [15] SALVATI, D. A., "Numerical Computation of Polynomials of Best Uniform Approximation to the Function $|x|$ ", *Master's Thesis* (Ohio State University, Columbus, Ohio, 1980).
 - [16] VARGA, R. S. and CARPENTER, A. J., "On the Bernstein Conjecture in approximation theory", *Constr. Approx.* **1** (1985), 333–348; *Ügyszintén megjelent oroszul: Mat. Sbornik* **129** (171) (1986), 535–548.
 - [17] VARGA, R. S., RUTTAN, A. and CARPENTER, A. J., "Numerical results on best uniform rational approximations to $|x|$ on $[-1, +1]$ ", *Mat. Sbornik* **182** (11) (1991), 1523–1541.
 - [18] VJACHESLAVOV, N. S., "On the least deviations of the function $\operatorname{sign} x$ and its primitives from the rational functions in the L_p -metrics, $0 < p \leq \infty$ ", *Mat. Sbornik* **103** (145) (1977), 24–36, (oroszul); *Angol fordítás: Math. USSR Sbornik* **32** (1977), 19–31.
 - [19] WHITTAKER, E. T. és WATSON, G. N., *A Course of Modern Analysis*, 4th ed. (Cambridge University Press, Cambridge, 1962).
 - [20] BERNSTEIN, S. N., "Sur la meilleure approximation de $|x|^p$ par des polynômes de degrés très élevés", *Bull. Acad. Sci. USSR, Cl. sci. math. nat.* **2** (1938), 181–190.
 - [21] BERNSTEIN, S. N., *Collected Works*, vol. II (Akad. Nauk SSSR, Moscow, 1954), 262–272, (oroszul).
 - [22] CARPENTER, A. J. and VARGA, R. S., "Some numerical results on best uniform polynomial approximation of x^α on $[0,1]$ ", *USSR-US Conference on Approximation Theory* 1991, (megjelenés alatt).
 - [23] STAHL, H., "Best uniform rational approximation of $|x|$ on $[-1, +1]$ ", *Mat. Sbornik*, (megjelenés alatt).
 - [24] VARGA, R. S. and CARPENTER, A. J., "On best uniform rational approximation of x^α on $[0,1]$ ", *Numerical Algorithms*, (megjelenés alatt).

2. FEJEZET

Az „1/9” sejtés és megoldása

2.1. Az „1/9” sejtés eredete

Az érdekesség kedvéért röviden ismertetjük az ún. „1/9” sejtést, valamint ennek Gonchar és Rakhmanov [6] által adott megoldását. Mint látni fogjuk, különösen meglepő az, hogy az e témával kapcsolatos kutatások oly különböző területeket hoznak egymással összefüggésbe, mint a parciális differenciálegyenletek numerikus megoldása, mátrixelmélet, racionális approximáció, nagypontosságú számítások és komplex potenciálmélet. Kiindulásként, az egyszerűség kedvéért tekintünk az alábbi másodrendű lineáris parabolikus parciális differenciálegyenlet numerikus megoldását:

$$(1.1) \quad \phi u_t(\mathbf{x}, t) = \sum_{i=1}^n (K_i(\mathbf{x}) u_{x_i}(\mathbf{x}, t))_{x_i} - \sigma(\mathbf{x}) u(\mathbf{x}, t) + S(\mathbf{x}),$$

tetszőleges $t > 0$ időben és bármely $\mathbf{x} \in \Omega$ mellett (ahol Ω az \mathbb{R}^n egy nyílt, összefüggő, korlátos részhalmaza), az

$$(1.2) \quad u(\mathbf{x}, t) = g(\mathbf{x}) \quad (\mathbf{x} \in \partial\Omega; t > 0),$$

peremérték-feltételekkel a $\partial\Omega$ -n (vagyis az Ω határán), valamint az

$$(1.3) \quad u(\mathbf{x}, 0) = h(\mathbf{x}) \quad (\mathbf{x} \in \Omega).$$

kezdetiérték-feltételekkel. Itt $S(\mathbf{x})$ és $h(\mathbf{x})$ adott, az $\bar{\Omega}$ -n (vagyis az Ω lezártján) definiált, megfelelően sima függvények, míg $g(\mathbf{x})$ egy adott, a $\partial\Omega$ -n definiált megfelelően sima függvény. Ezenkívül ϕ egy adott pozitív konstans, és a $K_i(\mathbf{x})$, valamint $\sigma(\mathbf{x})$ mennyiségekről feltételezzük, hogy pozitívak a $\bar{\Omega}$ -ban. E feltételezések fontos fizikai és mérnöki problémákat takarnak. Például a reaktor-fizikában egy reaktor valamely átlagenergiájában a neutronok időtől függő $u(\mathbf{x}, t)$ sűrűsége kielégíti (1.1)-et (egy diffúzió-elméleti approximáció esetén), aholis $1/\phi$ fizikailag reprezentálja ezen neutronok *átlagsebességét*, és ahol $K_i(\mathbf{x}) \equiv K(\mathbf{x})$ a *diffúziós együttható*, továbbá $\sigma(\mathbf{x})$ a *teljes elmozdítási keresztmetszet* a reaktor minden egyes pontjában. Hasonló alkalmazások említhetők az olajbányászattal kapcsolatban, egy homogén porózus közegben áramló, összenyomható folyadék esetén.

Egy \mathbb{R}^n -beli diszkrét partíciót, valamint egy megfelelő, $(2n + 1)$ pontból álló különbségi-hányados-approximációt alkalmazva a térbeli deriváltakra, az (1.1)–(1.3) egy szemi-diszkrét approximációja (vagyis egy diszkrét tér, de folytonos idő) azt eredményezi, hogy (lásd Varga [24, 253. oldal])

$$(1.4) \quad \begin{cases} \frac{d\mathbf{w}(t)}{dt} = -A\mathbf{w}(t) + \mathbf{r} & (t > 0), \\ \mathbf{w}(0) = \tilde{\mathbf{w}}, \end{cases}$$

ahol A egy ritka, valós, szimmetrikus, pozitív definit, $N \times N$ -es mátrix. Az (1.4) megoldása nyilván kifejezhető a

$$(1.5) \quad w(t) = A^{-1}r + \exp(-tA) \{ \tilde{w} - A^{-1}r \} \quad (t \geq 0),$$

alakban, ahol, mint szokásos, $\exp(-tA) := \sum_{k=0}^{\infty} (-tA)^k / k!$.

Jólismert (lásd [24, 8.3. pont]), hogy (1.5) megoldását általában az $\exp(-tA)$ racionális Padé approximációjával közelítik, és ezek, speciális esetekként kiadják a jólismert *haladó* és *retrográd differenciákat*, valamint a *Crank-Nicholson* módszert ilyen parabolikus problémák esetén. A következő paragrafusban minket nem annyira az $\exp(-tA)$ Padé-féle, mint inkább ennek racionális *Csebisev* approximációja érdekel majd. Ennek oka az, hogy az e^{-x} racionális Padé approximációja, mint egy $x = 0$ körül definiált *lokális* approximáció, általában rosszul approximálja az e^{-x} -et nagy x esetén, és ez (stabilitási és/vagy pontossági okokból) az idő-lépcsőkzzel kapcsolatban korlátozásokhoz vezetne. Ezzel ellentétben, az e^{-x} racionális *Csebisev* approximációja *globálisan* definiált a $[0, +\infty)$ intervallumon, és ezen approximációknál – mint azt látni fogjuk – nem lesz szükség az idő-lépcsőkz korlátozására.

2.2. Szemi-diszkrét Csebisev approximációk

Az (1.5) szemi-diszkrét *Csebisev* approximációinak meghatározásához tekintsük a következő approximációelméleti problémát. Ha π_m jelöli a legfeljebb m -edfokú valós polinomokat, és hasonlóképpen $\pi_{m,n}$ az összes valós $r_{m,n}(x) = p(x)/q(x)$ alakú racionális függvények halmaza, ahol $p \in \pi_m$ és $q \in \pi_n$, akkor legyen

$$(2.1) \quad \lambda_{m,n} = \lambda_{m,n}(e^{-x}) := \inf \{ \|e^{-x} - r_{m,n}(x)\|_{L_{\infty}[0,+\infty)} : r_{m,n} \in \pi_{m,n} \}.$$

E konstansokat az e^{-x} függvény $[0, +\infty)$ intervallumon vett egyenletes racionális *Csebisev konstansainak* nevezzük. Nyilvánvaló, hogy $\lambda_{m,n}$ véges, ha $0 \leq m \leq n$. Továbbá, ha (m, n) a nemnegatív $0 \leq m \leq n$ egészekből alkotott pár, akkor ismeretes (lásd Achieser [1, 55. oldal]), hogy az esetleges közös tényezőkkel való leosztás után, létezik egy egyértelmű $\hat{r}_{m,n} \in \pi_{m,n}$ úgy, hogy

$$(2.2) \quad \hat{r}_{m,n}(x) = \hat{p}_{m,n}(x) / \hat{q}_{m,n}(x) \quad (\hat{p}_{m,n} \in \pi_m; \hat{q}_{m,n} \in \pi_n),$$

ahol $\hat{q}_{m,n}(x) > 0$ a $[0, +\infty)$ intervallumon, és

$$(2.3) \quad \lambda_{m,n} = \|e^{-x} - \hat{r}_{m,n}(x)\|_{L_{\infty}[0,+\infty)}.$$

Ha $\hat{q}_{m,n}(x) := \sum_{j=0}^n c_j x^j$, akkor legyen $\hat{q}_{m,n}(tA) := \sum_{j=0}^n c_j (tA)^j$, s ez utóbbi az A $N \times N$ -es mátrix egy n -edfokú valós polinomja. Továbbá, mivel $\hat{q}_{m,n}(x) > 0$ a $[0, +\infty)$ intervallumon, és mivel A valós, szimmetrikus és pozitív definit, láthatjuk, hogy a

$\hat{q}_{m,n}(tA)$ mátrix is valós, szimmetrikus és pozitív definit (és így nem szinguláris) bármely $t \geq 0$ esetén.

Az (1.5)-tel analóg módon definiáljuk (mint az [23]-ban eredetileg is történt) az (1.5) $w(t)$ megoldásának (m, n) -ik *szemi-diszkrét Csebisev approximációját*, a

$$(2.4) \quad w_{m,n}(t) := A^{-1}r + (\hat{q}_{m,n}(tA))^{-1}(\hat{p}_{m,n}(tA))\{\tilde{w} - A^{-1}r\} \quad (t \geq 0),$$

mennyiséget, vagy ekvivalens módon legyen

$$(2.4') \quad \hat{q}_{m,n}(tA)w_{m,n}(t) = \hat{q}_{m,n}(tA)A^{-1}r + \hat{p}_{m,n}(tA)\{\tilde{w} - A^{-1}r\} \quad (t \geq 0).$$

Felbontva a $\hat{q}_{m,n}(x)$ polinomot az x valós lineáris, valamint kvadratikus polinomjainak szorzatára (e tényezők szükségszerűen pozitívak a $[0, +\infty)$ intervallumon), a (2.4') $w_{m,n}(t)$ megoldása elérhető úgy, hogy rekursíve megoldunk egy sorozat mátrix-problémát, ahol minden egyes probléma egy lineáris vagy kvadratikus szimmetrikus pozitív definit tA mátrix Gauss-féle eliminációval történő direkt inverzióját jelenti $t \geq 0$ esetén. (További részleteket illetően lásd Varga [25, 71. oldal].)

Mennyire jól approximálja a (2.4)-beli $w_{m,n}(t)$ az (1.5)-beli $w(t)$ mennyiséget? Az N -dimenziós \mathbb{C}^N komplex vektortérben ℓ_2 -vektornormát használva, ismeretes, hogy ha B egy $N \times N$ -es Hermite-féle mátrix a $\{\mu_j\}_{j=1}^N$ (valós) sajátértékekkel, akkor a B indukált ℓ_2 -operátor normája egyszerűen

$$\|B\|_2 = \max_{1 \leq j \leq N} \{|\mu_j|\}.$$

Következésképpen, ha $\{\alpha_j\}_{j=1}^N$ jelöli az (1.5)-beli $N \times N$ -es valós, szimmetrikus, pozitív definit A mátrix (pozitív) sajátértékeit, akkor

$$\|\exp(-tA) - \hat{r}_{m,n}(tA)\|_2 = \max_{1 \leq j \leq N} |e^{-t\alpha_j} - \hat{r}_{m,n}(t\alpha_j)|.$$

Viszont mivel $t\alpha_j \geq 0$ minden $1 \leq j \leq N$ és minden $t \geq 0$ esetén, a (2.3) maga után vonja, hogy

$$\|\exp(-tA) - \hat{r}_{m,n}(tA)\|_2 \leq \lambda_{m,n} \quad (t \geq 0).$$

Következésképpen, az (1.5) és (2.4) alapján

$$(2.5) \quad \|w(t) - w_{m,n}(t)\|_2 \leq \lambda_{m,n} \|\tilde{w} - A^{-1}r\|_2 \quad (t \geq 0).$$

Jegyezzük meg, hogy mivel (2.5) jobb oldala *független* a t -től, így egy hibakorlátot kaptunk a $w(t) - w_{m,n}(t)$ mennyiségre *minden* $t \geq 0$ esetén. A jólismert Padé módszerekkel szemben, amelyek a t nagyságát pontossági és/vagy stabilitási okokból korlátozzák, a szemi-diszkrét Csebisev módszer nagyon nagy t értékekre is használható. Nevezetesen, rögzítve az egészekből álló (m, n) számpárt úgy, hogy

$0 \leq m \leq n$, a (2.4) szemi-diszkrét Csebisev módszert úgy tekinthetjük, mint egy időlépéses eljárás *definícióját egyetlen tetszőleges pozitív idő-lépéssel*. (Neil Armstrong űrhajós egy kijelentését szabadon átalakítva, „kis lépés Padéval, nagy ugrás Csebisevvel!”) Mindazonáltal megjegyezzük, hogy a szemi-diszkrét Csebisev módszerek olyan időtől független együtthatójú lineáris problémákra *korlátozottak*, mint (1.1). Nem ilyen formájú fizikai problémák esetén az $\exp(-\Delta t A)$ függvény Padé-típusú approximációja szükségeltetik. A Padé approximációk kezdetiérték-feladatokra való újkéltű alkalmazásait illetően lásd Reusch et al [14].

2.3. Az e^{-x} Csebisev konstansai

A szemi-diszkrét Csebisev approximációk alkalmazhatósága (2.5) alapján *döntően* a (2.1)-beli $\lambda_{m,n}$ Csebisev-konstansok viselkedésétől függ, midőn $n \rightarrow \infty$. A (2.1)-ből nyilvánvaló, hogy

$$(3.1) \quad 0 < \lambda_{n,n} \leq \lambda_{n-1,n} \leq \cdots \leq \lambda_{0,n} \quad (n = 0, 1, \dots).$$

Elemi okoskodást használva, Cody, Meinardus és Varga [5] megmutatták, hogy ha $s_n(x) := \sum_{j=0}^n x^j/j!$ jelöli az e^x n -ik részletösszegét, akkor

$$(3.2) \quad \lim_{n \rightarrow \infty} \left\{ \left\| e^{-x} - \frac{1}{s_n(x)} \right\|_{L_\infty[0, +\infty)} \right\}^{1/n} \leq \frac{1}{2}.$$

Mivel $1/s_n(x) \in \pi_{0,n}$, a (3.2)-ből és a $\lambda_{0,n}$ (2.1)-beli definíciójából közvetlenül következik, hogy

$$(3.3) \quad \overline{\lim}_{n \rightarrow \infty} \lambda_{0,n}^{1/n} \leq \frac{1}{2}.$$

Innen, ha $\{m(n)\}_{n=0}^\infty$ a nemnegatív egészek egy tetszőleges olyan sorozata, hogy $0 \leq m(n) \leq n$, akkor (3.3) és a (3.1)-beli egyenlőtlenségek maguk után vonják, hogy

$$(3.4) \quad \overline{\lim}_{n \rightarrow \infty} (\lambda_{m(n),n})^{1/n} \leq \frac{1}{2}.$$

A (3.4)-beli $1/2$ felső korlát némileg javítható egy ún. eltolási módszer segítségével:

1. TÉTEL ([5]). Legyen $\{m(n)\}_{n=0}^\infty$ nemnegatív számok egy tetszőleges sorozata úgy, hogy $0 \leq m(n) \leq n$ bármely $n \geq 0$ esetén. Ekkor

$$(3.5) \quad \frac{1}{6} \leq \overline{\lim}_{n \rightarrow \infty} (\lambda_{m(n),n})^{1/n} \leq \frac{1}{2,29878}.$$

A (3.5)-beli eredmény nyilván azt jelenti, hogy az e^{-x} függvény $\lambda_{m,n}$ Csebisev-konstansai a $[0, +\infty)$ intervallumon *geometriai sebességgel* konvergálnak zérushoz.

n	$\lambda_{n,n}$	$1/\lambda_{n,n}^{1/n}$	n	$\lambda_{n,n}$	$1/\lambda_{n,n}^{1/n}$	n	$\lambda_{n,n}$	$1/\lambda_{n,n}^{1/n}$
0	5,000 (-1)	—	5	9,346 (-6)	10,14	10	1,361 (-10)	9,696
1	6,685 (-2)	14,96	6	1,008 (-6)	9,987	11	1,466 (-11)	9,658
2	7,359 (-3)	11,66	7	1,087 (-7)	9,882	12	1,579 (-12)	9,626
3	7,994 (-4)	10,77	8	1,172 (-8)	9,804	13	1,701 (-13)	9,600
4	8,653 (-5)	10,37	9	1,263 (-9)	9,744	14	1,832 (-14)	9,577

2.1. Táblázat

Mivel (3.1) alapján a $\{\lambda_{n,n}\}_{n=0}^{\infty}$ sorozat szükségszerűen a *leggyorsabban* konvergál zérushoz, Cody, Meinardus és Varga [5] 1969-ben kiszámították (a Remez algoritmus segítségével) a $\{\lambda_{n,n}\}_{n=0}^{14}$ első becsléseit, melyeket a 2.1. Táblázatban adtunk meg.

Így a $\lambda_{n,n}$ geometriai nagyságrendű konvergenciája *lényegesen* jobbnak tűnik, mint a (3.5)-beli felső korlátra vonatkozó becslés. Később, 1973-ban Schönhage [18] bebizonyította, hogy

$$\frac{1}{6\sqrt{(4n+4)\log 3 + 2\log 2}} \leq 3^n \lambda_{0,n} \leq \sqrt{2} \quad (n = 0, 1, \dots),$$

és így valójában

$$(3.6) \quad \lim_{n \rightarrow \infty} \lambda_{0,n}^{1/n} = \frac{1}{3}.$$

Ilymódon, mivel a $\lambda_{n,n}$ -et meghatározó $\hat{r}_{n,n}(x)$ racionális függvényben szereplő együtthatók száma lényegében *kétszerese* a $\lambda_{0,n}$ -et meghatározó $\hat{r}_{0,n}(x)$ együtthatói számának, a (3.6) alatti Schönhage-féle eredmény és a 2.1. Táblázat számai (enyhén) azt sugallták, hogy 1977-ben kimondjuk az alábbi sejtést:

$$(3.7) \quad \lim_{n \rightarrow \infty} \lambda_{n,n}^{1/n} \stackrel{?}{=} \frac{1}{9}, \quad (\text{Saff és Varga [17]}).$$

Ezután a 2.1. Táblázatbeli, $\{\lambda_{n,n}\}_{n=0}^{14}$ értékekre vonatkozó, 1969-ből származó becslések *javítása* következett 1984-ben Carpenter, Ruttan és Varga [4] által, akik Richard Brent [2] többszörös pontosságú programjával 230 értékes jegyet kaptak. Csakúgy, mint az 1. fejezetben, a (második) Remez algoritmus használatával e számítások mintegy 200 értékes jegyre megadták a $\{\lambda_{n,n}\}_{n=0}^{30}$ Csebisev-konstansokat. E számokat a 2.2. Táblázatban tüntettük fel, 50 értékes jegyre kerekítve. Ezekből kiszámítottuk [4] a $\{\lambda_{n-1,n-1}/\lambda_{n,n}\}_{n=1}^{30}$ hányadosokat, melyeket a 2.3. Táblázatban adtuk meg.

A 2.3. Táblázat utolsó tizenegy elemét Richardson-féle extrapolációval számítottuk ki ($x_n = 1/n^2$), pontosan ugyanúgy, mint az 1. fejezetbeli numerikus számításoknál, a Bernstein-féle β konstans. Ezen extrapolációkat a 2.4–2.9. Táblázatokban adtuk meg.

n	$\lambda_{n-1,n-1}/\lambda_{n,n}$
1	7, 4815532397221509829356536616817047817627984227696 (+00)
2	9, 0819455991000169708588696090116053013062015321681 (+00)
3	9, 2054646248528427537813883233351970910096341895810 (+00)
4	9, 2390013695637342229492228910895668594203903511860 (+00)
5	9, 2579780201008948071386386966176824253418736773867 (+00)
6	9, 2673633886078728002406169193047563695298406988185 (+00)
7	9, 2731650684028757880126091410302398681948193463184 (+00)
8	9, 2768895688704833336324198589706052242724880476172 (+00)
9	9, 2794442071765347804120940269531440575883814379160 (+00)
10	9, 2812683495309755120464682831533111454037182839070 (+00)
11	9, 2826170054814049413318434721810547537510453765262 (+00)
12	9, 2836420758101343365763572286526847004054823670187 (+00)
13	9, 2844394306651793615709775312796518943731046644594 (+00)
14	9, 2850718606898552364565090058954105627072638256542 (+00)
15	9, 2855819149043751995853520164465592457350862957931 (+00)
16	9, 2859992519340952301903430112062879793060351437360 (+00)
17	9, 2863450591648612312069660869905305380460379908323 (+00)
18	9, 2866347991400778934888938646138283512490395348046 (+00)
19	9, 2868799705918004397301050888494834728534387767879 (+00)
20	9, 2870892682832631479754585814309629455011369059197 (+00)
21	9, 287269365333554026168814849193491692724731187991 (+00)
22	9, 2874254522088778823625744929007529809804239100012 (+00)
23	9, 2875616152014957847981034264879772236314002103896 (+00)
24	9, 2876811067903443960632445539185855003199200785395 (+00)
25	9, 2877865418013514399321174449489929772550616085232 (+00)
26	9, 2878800417598657237157599122881662160311796009207 (+00)
27	9, 2879633425048853224353278979631458037287458900494 (+00)
28	9, 2880378753756707950994314690770008799076337263447 (+00)
29	9, 2881048291364217038868850048114151489211447279412 (+00)
30	9, 2881651976905816378400677087169532012794506558118 (+00)

2.9. Táblázat $\{\lambda_{n-1,n-1}/\lambda_{n,n}\}_{n=1}^{30}$

zöképpen. Legyen

$$(3.9) \quad \exp[(x-1)/(x+1)] = \sum_{k=0}^{\infty} c_k T_k(x) \quad (x \in [-1, +1])$$

az $\exp[(x-1)/(x+1)]$ függvény $[-1, +1]$ intervallumon vett Csebisev-sorfejtése, ahol

$$(3.10) \quad c_k := \frac{2}{\pi} \int_{-1}^{+1} \exp[(x-1)/(x+1)] T_k(x) dx / \sqrt{1-x^2} \quad (k = 0, 1, \dots),$$

<i>Első Richardson-féle extrapoláció</i>	
9.2890264097244993503380805433426294011879483664028	(+00)
9.289026250164817732777051096402664917526634197715	(+00)
9.2890261238332083354380146232927890779219008634554	(+00)
9.2890260227584489781751521371269212102821968923977	(+00)
9.2890259411144138331662151027758237265334600018012	(+00)
9.2890258745847956720447117179152892402483118606939	(+00)
9.2890258199319277513867988095911873373807234645966	(+00)
9.2890257746993546055018224389315527078059834110588	(+00)
9.2890257370035920984371231103514289542999627148124	(+00)
9.2890257053863190014778076067603006933698792208821	(+00)

2.4. Táblázat

<i>Második Richardson-féle extrapoláció</i>	
9.2890254903568100301092506634385383515846398643843	(+00)
9.2890254907395748783152475269378170381563748710076	(+00)
9.2890254910173236638792234055586598196559281750941	(+00)
9.2890254912217618361898849864036303004596910942434	(+00)
9.2890254913741950639848522211186105992458585679159	(+00)
9.2890254914892118820483798893742287864937739688197	(+00)
9.2890254915769523219995629336916174338011290018776	(+00)
9.2890254916445664944177855162897305944259530044229	(+00)
9.2890254916971628978907097204202614126010390550083	(+00)

2.5. Táblázat

<i>Az „1/9” sejtés és megoldása</i>	
9.2890254919264426246904315998037974616346229535605	(+00)
9.2890254919246363633882112757200795725544689684432	(+00)
9.2890254919235212361782552355452310998144517528836	(+00)
9.2890254919227472919000744467255805180069646466419	(+00)
9.2890254919222163735816605228071443149565143016341	(+00)
9.2890254919218439884743390512914155757325560186146	(+00)
9.2890254919215797099009277334795760282585652815175	(+00)
9.2890254919213896705910708011872612164000901127673	(+00)

2.6. Táblázat

és a (3.9)-beli összegben a vessző azt jelenti, hogy c_0 helyett $c_0/2$ szerepel. A $H : [c_{i+j-1}]_{i,j=1}^{\infty}$ végtelen Hankel-mátrixszal kapcsolatban legyen

$$\sigma_n := \text{a } H \text{ } n\text{-edik szinguláris értéke} \quad (\text{ahol } \sigma_1 \geq \sigma_2 \geq \dots).$$

<i>Negyedik Richardson-féle extrapoláció</i>	
9, 2890254919205312240649832664389025519177553659037	(+00)
9, 2890254919208485671587410305609475319973452742869	(+00)
9, 2890254919207963073654937082427113430339242329911	(+00)
9, 2890254919208120946294556940428805578880731390886	(+00)
9, 2890254919208127681771411301709359901892869271757	(+00)
9, 2890254919208150149547144296258736339472494913060	(+00)
9, 2890254919208161591023954162336683020414775499319	(+00)

2.7. Táblázat

<i>ötödik Richardson-féle extrapoláció</i>	
9, 2890254919214127326587548334445830521388384447460	(+00)
9, 2890254919206982368598678821050851331068234703891	(+00)
9, 2890254919208432825305071272154597415183101617465	(+00)
9, 2890254919208141654584179760993333379827853346193	(+00)
9, 2890254919208198985165341295732097954739907627740	(+00)
9, 2890254919208187594380340221604743658919958649908	(+00)

2.8. Táblázat

<i>Hatodik Richardson-féle extrapoláció</i>	
9, 2890254919196627357020607062507403229154974205964	(+00)
9, 2890254919210653837136734712907208606483991578876	(+00)
9, 2890254919207671899154474789653160736121386135206	(+00)
9, 2890254919208296189900708128670580198741435560234	(+00)
9, 2890254919208167344095893867600558244128938244875	(+00)

2.9. Táblázat

A [22]-ben azt sejtették, hogy

$$\lambda_{n,n} \stackrel{?}{\sim} \sigma_n \quad (\text{midőn } n \rightarrow \infty),$$

és a σ_n numerikus becsléseire alapozva, Trefethen és Gutknecht [22] azt gondolták, hogy

$$(3.11) \quad \lim_{n \rightarrow \infty} \lambda_{n,n}^{1/n} \stackrel{?}{=} \frac{1}{9,28903}.$$

A (3.8) és (3.11) nagyon hasonló numerikus becslései, melyek teljesen más numerikus módszerekken alapultak, erősen azt sugallták, hogy a (3.7) sejtés *nem igaz*.

Az „1/9” sejtéssel nagyon sok kutató foglalkozott; az 1982 évvel bezárólag ezek összefoglalása megtalálható Varga [26] munkájában. E kutatások több különböző

irányban folytatódtak. Először is alsó korlát becsléseket kerestek Λ_1 -re, valamint felső korlát becsléseket Λ_2 -re, ahol (lásd (2.1))

$$(3.12) \quad \Lambda_1 := \lim_{n \rightarrow \infty} \lambda_{n,n}^{1/n}(e^{-x}) \quad \text{és} \quad \Lambda_2 := \overline{\lim}_{n \rightarrow \infty} \lambda_{n,n}^{1/n}(e^{-x}).$$

E mennyiségek tehát a *specifikus* e^{-x} függvénynek a $[0, +\infty)$ intervallumon racionális függvényekkel történő megközelítési hibájának geometriai konvergencia-sebességét mérik. A (3.12)-beli Λ_1 -re vonatkozó legjobb ilyen alsó korlátot Schönhage [19] találta 1982-ben, aki megmutatta, hogy

$$(3.13) \quad \frac{1}{13,928} < \Lambda_1 := \lim_{n \rightarrow \infty} \lambda_{n,n}^{1/n}(e^{-x}),$$

míg a (3.9)-beli Λ_2 -re a legjobb felső korlátot Opitz és Scherer [12] határozta meg 1985-ben, akik megmutatták, hogy

$$(3.14) \quad \overline{\lim}_{n \rightarrow \infty} \lambda_{n,n}^{1/n}(e^{-x}) =: \Lambda_2 < \frac{1}{9,037}.$$

Ezen eredmény természetesen *bizonyította*, hogy a (3.7)-beli "1/9" sejtés *nem igaz*. Egyébként (3.14) azt is bizonyítja, hogy a $\{\lambda_{n,n}(e^{-x})\}_{n=0}^{\infty}$ konstansok zérushoz való geometriai konvergenciájának *foka* ténylegesen *jobb*, mint 1/9!

Egy másik irány: természetes azt kérdezni, hogy léteznek-e az e^x -től különböző $f(x) (\neq 0)$ folytonos függvények a $[0, +\infty)$ intervallumon úgy, hogy az $1/f(x)$ -re vonatkozó Csebisev-konstansok (lásd (2.1)), vagyis a

$$(3.15) \quad \lambda_{m,n}(1/f) := \inf \{ \|1/f(x) - r_{m,n}(x)\|_{L_{\infty}[0,\infty)} : r_{m,n} \in \pi_{m,n} \},$$

mennyiségek geometriailag konvergálnak zérushoz; másszóval létezik-e a nemnegatív egészeknek egy $\{m(n)\}_{n=0}^{\infty}$ sorozata úgy, hogy $0 \leq m(n) \leq n$ bármely $n \geq 0$ esetén, és emellett

$$(3.16) \quad \overline{\lim}_{n \rightarrow \infty} (\lambda_{m(n),n}(1/f))^{1/n} < 1.$$

Ahhoz, hogy az ezirányú eredményeket leírassuk, egy jelölésre van szükségünk. Adott $r > 0$ és $s > 1$ mellett, legyen

$$(3.17) \quad \mathcal{E}(r, s) := \left\{ z = x + iy : \frac{(x - \frac{r}{2})^2}{[\frac{r}{4}(s + \frac{1}{s})]^2} + \frac{y^2}{[\frac{r}{4}(s - \frac{1}{s})]^2} \leq 1 \right\}$$

egy zárt ellipszis a komplex síkon, és ha $g(z)$ az ezen $\mathcal{E}(r, s)$ ellipszisen definiált függvény, akkor legyen

$$(3.18) \quad \tilde{M}_g(r, s) := \sup \{ |g(z)| : z \in \mathcal{E}(r, s) \}.$$

E jelöléssel a következő eredményt fogalmazhatjuk meg:

2. TÉTEL (Meinardus et al [9]). Legyen $f(x)$ egy valós, folytonos függvény ($\neq 0$) a $[0, +\infty)$ intervallumon, és tegyük fel, hogy létezik a valós polinomoknak egy $\{p_n(x)\}_{n=0}^\infty$ sorozata úgy, hogy $p_n \in \pi_n$ bármely $n \geq 0$ esetén, valamint egy valós $q > 1$ úgy, hogy

$$(3.19) \quad \overline{\lim}_{n \rightarrow \infty} \{ \|1/p_n - 1/f\|_{L_\infty[0, \infty)} \}^{1/n} \leq \frac{1}{q} < 1.$$

Ekkor $f(x)$ kiterjeszthető egy $F(z)$ egész függvénné úgy, hogy $F(x) = f(x)$ bármely $x \geq 0$ esetén, és $F(z)$ véges ρ rendű, vagyis ha $M_F(r) := \max_{|z|=r} |F(z)|$, akkor

$$(3.20) \quad \overline{\lim}_{r \rightarrow \infty} \frac{\log \log M_F(r)}{\log r} = \rho, \text{ ahol } 0 \leq \rho < \infty.$$

Továbbá bármely $s > 1$ esetén léteznek K, θ és r_0 konstansok (mind függetlenek s és q értékétől) úgy, hogy (lásd (3.18))

$$(3.21) \quad \tilde{M}_F(r, s) \leq K (\|f\|_{L_\infty[0, r]})^\theta \quad (\text{bármely } r \geq r_0 \text{ esetén}).$$

Fordítva, legyen $f(z) = \sum_{k=0}^\infty a_k z^k$ egy valós egész függvény úgy, hogy $a_0 > 0$ és $a_k \geq 0$ bármely $k \geq 1$ esetén. Ha léteznek valós $A > 0$, $s > 1$, $\theta > 0$ és $r_0 > 0$ számok úgy, hogy (lásd (3.18) és (3.21))

$$(3.22) \quad \tilde{M}_f(r, s) \leq A (\|f\|_{L_\infty[0, r]})^\theta \quad (\text{bármely } r \geq r_0 \text{ esetén}),$$

akkor létezik a valós polinomok egy $\{p_n(x)\}_{n=0}^\infty$ sorozata úgy, hogy $p_n \in \pi_n$ bármely $n \geq 0$ és valós $q \geq s^{1/(1+\theta)} > 1$ mellett

$$(3.23) \quad \overline{\lim}_{n \rightarrow \infty} \{ \|1/p_n - 1/f\|_{L_\infty[0, \infty)} \}^{1/n} = \frac{1}{q} < 1.$$

Következésképpen, ha $m(n) = 0$ bármely $n \geq 0$ esetén (lásd (3.15)), akkor

$$(3.24) \quad \overline{\lim}_{n \rightarrow \infty} (\lambda_{0,n}(1/f))^{1/n} \leq \frac{1}{q} < 1.$$

Érdekes, hogy az ellipsziseken vett növekedési ráta fellép a 2. Tételben a geometriai rendű konvergencia szükséges és elégséges feltételeiben (lásd (3.21)-(3.22)). Mint láthatjuk, van egy kis hézag a 2. Tételben a szükséges és elégséges feltételek között, s ezt még nem sikerült kiküszöbölni.

Egy másik, a problémával kapcsolatban álló, de tőle különböző irányt jelent az e^{-x} -hez tartozó egyenletes racionális Csebisev konstansok zéróhoz tartó geometriai rendű konvergenciájának kérdése, csakúgy, mint (3.2)-ben, de most a $[0, +\infty)$

intervallumtól különböző *tartományokban*, mint pl. a komplex síkon a $[0, +\infty)$ sugár körüli szimmetrikus *szektorokban*. Ilyen eredmények megfogalmazása céljából, a nemnegatív egészek egy (ν, n) párjához hozzárendeljük az e^{-z} (ν, n) -ik $R_{\nu, n}(x)$ racionális Padé approximációját (a $\pi_{\nu, n}$ -ben), amelyet explicite (lásd Perron [13, 433. oldal]) az

$$(3.25) \quad R_{\nu, n}(x) = Q_{\nu, n}(x)/P_{\nu, n}(x)$$

alakban adhatunk meg, ahol

$$(3.26) \quad Q_{\nu, n}(x) := \sum_{j=0}^{\nu} \frac{(n+\nu-j)! \nu! (-z)^j}{(n+\nu)! j! (\nu-j)!}, \text{ és } P_{\nu, n}(z) := \sum_{j=0}^n \frac{(n+\nu-j)! n! z^j}{(n+\nu)! j! (n-j)!}.$$

Továbbá bármely valós θ esetén, amelyre $0 \leq \theta < \pi$, definiáljuk a

$$(3.27) \quad S(\theta) := \{z \in \mathbb{C} : |\arg z| \leq \theta\}$$

szimmetrikus szektort. Ezenkívül bármely, az $S(\theta)$ -n definiált $h(z)$ függvényre legyen

$$(3.28) \quad \|h\|_{L_{\infty}(S(\theta))} := \sup\{|h(z)| : z \in S(\theta)\}.$$

E jelöléssel kimondhatjuk az alábbi tételt.

3. TÉTEL (Saff, Varga és Ni [15]). *Tegyük fel, hogy a (3.25)–(3.26) által adott $\{R_{\nu(n), n}(z)\}_{n=0}^{\infty}$ racionális Padé approximációk sorozata kielégíti a következő feltételt:*

$$(3.29) \quad \lim_{n \rightarrow \infty} \frac{\nu(n)}{n} = \sigma, \text{ ahol } 0 < \sigma < 1,$$

és legyen

$$(3.30) \quad \theta_0 := \cos^{-1} \left(\frac{1-\sigma}{1+\sigma} \right), \text{ és } g(\sigma) := \sigma^{\sigma} (1-\sigma)^{1-\sigma} / 2^{1-\sigma}.$$

Ekkor bármely olyan θ esetén, amely kielégíti a

$$(3.31) \quad 0 < \theta < 4 \tan^{-1} \left\{ \left(\frac{1 - \sqrt{g(\sigma)}}{1 + \sqrt{g(\sigma)}} \right) \tan \left(\frac{\theta_0}{4} \right) \right\}$$

feltételt, teljesül, hogy

$$(3.32) \quad \overline{\lim}_{n \rightarrow \infty} \left\{ \|e^{-z} - R_{\nu(n), n}(z)\|_{L_{\infty}(S(\theta))} \right\}^{1/n} \leq g(\sigma) \left\{ \frac{\sin[1/4(\theta_0 + \theta)]}{\sin[1/4(\theta_0 - \theta)]} \right\} < 1.$$

Az $S(\theta)$ szektorban az e^{-x} -re vonatkozó, nem Padé-jellegű, de hasonlóan geometriai sebességű racionális approximáció található Saff, Schönhage és Varga [16] munkájában.

2.4. Gonchar és Rakhmanov eredményei

Az elmélet egy szép és mély továbbfejlesztéseként Gonchar és Rakhmanov [6], a komplex számsík potenciálméletének alkalmazásával *pontos* megoldást találtak az "1/9" sejtésre. Sajnos, e módszer nem írható le megfelelő módon egy pár oldalon. Ugyanezen témakör fejlődésében fontos szerepet játszott Nuttallnak [11] a kompakt Riemann felületeken vett Abel integrálok elméletére alapozott lokális racionális approximációs eredménye, valamint Stahlnek [20] a többpontos racionális Padé approximáció aszimptotikus viselkedésére kidolgozott tételei. Ezen eredmények áttekintésére vonatkozóan lásd Stahl [21] munkáját.

Gonchar és Rakhmanov [6] eredményeinek egy speciális esete a következő tétel.

4. TÉTEL (Gonchar és Rakhmanov [6]). A (2.1) alatt definiált $\lambda_{n,n}(e^{-x})$ mennyiséggel kapcsolatban létezik egy pozitív Λ szám úgy, hogy $0 < \Lambda < 1$ és

$$(4.1) \quad \lim_{n \rightarrow \infty} (\lambda_{n,n}(e^{-x}))^{1/n} = \Lambda.$$

Ez az eredmény természetesen azt jelenti, hogy a (3.12) alatti Λ_1 és Λ_2 mennyiségek *egyenlőek*. De hogy numerikusan mi ez a Λ szám, és hogy hogyan jellemezhető, az nagyon érdekes!

Kiderült, hogy korábban, 1986-ban Magnus [8] helyesen látta meg (pontos bizonyítás nélkül), hogy

$$(4.2) \quad \Lambda = \exp(-\pi K'/K) = \frac{1}{9,28902\ 54919\ 20818\ 91875\ 54494\ 35951 \dots},$$

ahol K és K' a k és $k' := \sqrt{1-k^2}$ modulusokra vonatkozó teljes elliptikus integrálok, azon pontban kiszámítva, ahol $K = 2E$, E lévén a teljes másodfajú elliptikus integrál. Másrészt 1986 augusztusában, a Berkeley-i Nemzetközi Matematikai Kongresszuson Gonchar bejelentette az alábbi eredményt:

5. TÉTEL (Gonchar és Rakhmanov [6]). A (4.1) alatti Λ szám a következő számelméleti jellemzéssel bír. Legyen

$$(4.3) \quad f(z) := \sum_{j=1}^{\infty} a_j z^j,$$

ahol

$$(4.4) \quad a_j := \left| \sum_{d|j} (-1)^d d \right| \quad (j = 1, 2, \dots),$$

úgyhogy $f(z)$ analitikus a $|z| < 1$ körben. Ekkor Λ az

$$(4.5) \quad f(\Lambda) = \frac{1}{8}$$

egyenlet egyetlen pozitív gyöke.

A Newton módszert használva, Carpenter [3] (4.5)-ből nagy pontossággal kiszámította a Λ -t, és 101 értékes jegyre az $1/\Lambda$ értéke

$$(4.6) \quad \frac{1}{\Lambda} = 9,28902\ 54919\ 20818\ 91875\ 54494\ 35951\ 74506\ 10316\ 94867\ 75012 \\ 44082\ 39700\ 61421\ 72937\ 52472\ 86507\ 07052\ 41587\ 06142\ 47144 \dots,$$

ami megerősíti a (3.8) és (3.11) alatti numerikus approximációkat.

Egy igazán érdekes fejleményként 1986 végén Magnus azt írta Goncharnak, hogy a (4.1)-beli Λ úgyszintén egyértelmű pozitív (egynél kisebb) megoldása a

$$(4.7) \quad \sum_{n=0}^{\infty} (2n+1)^2 (-\Lambda)^{n(n+1)/2} = 0,$$

egyenletnek, ami ekvivalens a (4.5) alatti megfogalmazással, továbbá hogy pontosan száz évvel korábban, 1886-ban Halphen [7] (4.7) segítségével kiszámította a Λ értékét hat értékes jegyre! (Halphen úgy jutott a (4.7) egyenlethez, hogy a theta-függvények variációit tanulmányozta.) Így méltányos, hogy az "1/9" konstanst Halphen konstansnak nevezzük!

Végül, hogy a 2.3. pontbeli 2. és 3. Tétellel való kapcsolatra rámutassunk, megjegyezzük (lásd [6, 329. oldal], hogy Gonchar és Rakhmanov [6] fő eredménye, vagyis az 1. Tétel, úgyszintén lehetővé teszi, hogy i) meghatározzuk az $e^{-p(x)}$ legjobb egyenletes racionális approximációjának mértékét a $[0, +\infty)$ intervallumon, ahol $p(x)$ egy tetszőleges, pozitív főegyütthatójú valós polinom, és ii) meghatározzuk az e^{-x} legjobb egyenletes racionális approximációját a (3.27)-beli $S(\theta) := \{z : |\arg z| \leq \theta\}$ szektorban, $\theta < \pi/2$ esetén. Az utóbbi esetben az eredményt a következőképpen fogalmazhatjuk meg:

6. TÉTEL ([6]). Bármely $n = 0, 1, \dots$ és $0 \leq \theta < \pi/2$ esetén legyen

$$(4.8) \quad \lambda_{n,n}(e^{-x}; S(\theta)) := \inf \{ \|e^{-x} - r_{n,n}(x)\|_{L_{\infty}(S(\theta))} : r_{n,n} \in \pi_{n,n} \}.$$

Ekkor a

$$(4.9) \quad \lim_{n \rightarrow \infty} (\lambda_{n,n}(e^{-x}; S(\theta)))^{1/n} = v_{\theta} \in (0, 1)$$

határérték létezik.

A (4.9)-beli v_{θ} értéke [6]-ban explicite adott, a theta-függvényeket tartalmazó integrál formájában. Nevezetesen

$$(4.10) \quad v_{\pi/4} \doteq \frac{1}{4.42}.$$

IRODALOM

- [1] ACHIESER, N. I., *Theory of Approximation* (Frederick Ungar Publishing Co., New York, 1956).
- [2] BRENT, RICHARD P., „A FORTRAN multiple-precision arithmetic package”, *Assoc. Comput. Mach. Trans. Math. Software* 4 (1978), 57–70.
- [3] CARPENTER, A. J., *Some theoretical and computational aspects of approximation theory*, Ph.D. tézis (The University of Leeds, Leeds, England, 1988).
- [4] CARPENTER, A. J., RUTTAN, A. és VARGA R. S., „Extended numerical computations on the „1/9” Conjecture in rational approximation theory”, *Rational Approximation and Interpolation*, (P.R. Graves-Morris, E.B. Saff, és R.S. Varga szerkesztők), *Lecture Notes in Mathematics* 1105 (Springer-Verlag, Heidelberg, 1984), 383–411.
- [5] CODY, W. J., MEINARDUS, G. és VARGA, R. S., „Chebyshev rational approximation to e^{-x} on $[0, +\infty)$ and applications to heat-conduction problems”, *J. Approx. Theory* 2 (1969), 50–65.
- [6] GONCHAR, A. A. és RAKHMANOV, E. A., „Equilibrium distribution and the degree of rational approximation of analytic functions”, *Mat. Sbornik* 134 176 (1987), 306–352, (oroszul); *Math. USSR Sbornik* 62 (1989), 305–348, (angolul).
- [7] HALPHEN, G.-H., *Traité des fonctions elliptiques et de leurs applications* (Gauthier-Villars, Paris, 1886).
- [8] MAGNUS, A. P., „CFGT determination of Varga's constant „1/9””, Preprint B-1348, *Inst. Math. Katholice Univ. Leuven, Louvain*, (1986).
- [9] MEINARDUS, G., REDDY, A. R., TAYLOR, G. D. és VARGA, R. S., „Converse theorems and extensions in Chebyshev rational approximation to certain entire functions in $[0, +\infty)$ ”, *Trans. Amer. Math. Soc.* 170 (1972), 171–186.
- [10] MEINARDUS, G. és VARGA, R. S., „Chebyshev rational approximations to certain entire functions in $[0, +\infty)$ ”, *J. Approx. Theory* 3 (1970), 300–309.
- [11] NUTTAL, J., „Asymptotics of diagonal Hermite–Padé polynomials”, *J. Approx. Theory* 42 (1984), 299–386.
- [12] OPITZ, H.-U. és SCHERER, K., „On the rational approximation of e^{-x} on $[0, \infty)$ ”, *Constr. Approx.* 1 (1985), 195–216.
- [13] PERRON, O., *Die Lehre von den Kettenbrüchen* (Teubner, Leipzig, 1929).
- [14] REUSCH, M. F., RATZAN, L., POMPHREY, N. és PARK, W., „Diagonal Padé approximations for initial value problems”, *SIAM J. Sci. Stat. Comput.* 9 (1988), 829–838.
- [15] SAFF, E. B., VARGA, R. S. és NI, W.-C., „Geometric convergence of rational approximations to e^{-x} in infinite sectors”, *Numer. Math.* 26 (1976), 211–225.
- [16] SAFF, E. B., SCHÖNHAGE, A. és VARGA, R. S., „Geometric convergence to e^{-x} by rational functions with real poles”, *Numer. Math.* 25 (1976), 307–322.
- [17] SAFF, E. B. és VARGA, R. S., „Some open questions concerning polynomials and rational functions”, *Padé and Rational Approximation*, (E.B. Saff és R.S. Varga szerkesztők) (Academic Press, Inc., New York, 1977), 483–488.
- [18] SCHÖNHAGE, A., „Zur rationalen Approximierbarkeit von e^{-x} über $[0, \infty)$ ”, *J. Approx. Theory* 7 (1973), 395–398.
- [19] SCHÖNHAGE, A., „Rational approximations to e^{-x} and related L^2 -problems”, *SIAM J. Numer. Anal.* 19 (1982), 1067–1082.
- [20] STAHL, H., „Orthogonal polynomials with complex-valued weight function I, II”, *Constr. Approx.* 2 (1986), 225–240, 241–251.
- [21] STAHL, H., „General convergence results for rational approximants”, *Approximation VI*, volume II (C.K. Chui, L.L. Schumaker, és J.D. Ward szerkesztők) (Academic Press, Inc., Boston, 1989), 605–634.
- [22] TREFETHEN, L. N. és GUTKNECHT, M. H., „The Carathéodory–Fejér method for real rational approximation”, *SIAM J. Numer. Anal.* 20 (1983), 420–436.
- [23] VARGA, R. S., „On higher order stable implicit methods for solving parabolic partial differential equations”, *J. Math. and Phys.* XL (1961), 220–231.
- [24] VARGA, R. S., *Matrix Iterative Analysis* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1962).

- [25] VARGA, R. S., *Functional Analysis and Approximation Theory in Numerical Analysis*, CBMS-NSF Regional Conference Series in Applied Math., vol. 3 (Soc. Ind. Appl. Math., Philadelphia, 1971).
- [26] VARGA, R. S., *Topics in Polynomial and Rational Interpolation and Approximation* (University of Montreal Press, Montreal, 1982).

3. FEJEZET

A Riemann hipotézis elméleti és számítástechnikai szempontból

3.1. Bevezetés

A modern matematika egyik leghíresebb megoldatlan problémája kétségtelenül az ún. *Riemann hipotézis*. Ennek felidézése céljából definiáljuk a

$$(1.1) \quad \zeta(z) := \sum_{n=1}^{\infty} \frac{1}{n^z}$$

Riemann-féle zeta-függvényt. Az (1.1) Dirichlet-sora konvergál bármely $Re\ z > 1$ esetén, mégpedig egyenletesen a $Re\ z \geq \sigma$ bármely részhalmazán, hacsak $\sigma > 1$ egy rögzített szám, s így $\zeta(z)$ analitikus a $Re\ z > 1$ tartományban. Úgyszintén ismeretes, hogy $\zeta(z)$ előállítható a

$$(1.2) \quad \zeta(z) = \prod_{p \text{ prím}} \left(1 - \frac{1}{p^z}\right)^{-1}$$

alakban, ami azt jelzi, hogy $\zeta(z)$ alapvető fontosságú az analitikus számelméletben. Mivel ismert, hogy az (1.2) végtelen szorzat reprezentáció bármely rögzített $\sigma > 1$ esetén abszolút konvergens a $Re\ z \geq \sigma$ zárt félsíkban, és mivel ezen végtelen szorzat egyetlen tényezője sem tűnik el a $Re\ z \geq \sigma$ tartományban, ezért $\zeta(z)$ nem zérus a $Re\ z > 1$ nyílt félsíkban. Az analitikus folytatás elvéből kifolyólag (lásd Titchmarsh [32, 13. oldal]), $\zeta(z)$ analitikus bármely komplex z pontban, kivéve $z = 1$ esetén, ami egy egyszeres pólusa a $\zeta(z)$ függvénynek 1 reziduummal; továbbá $\zeta(z)$ kielégíti a

$$(1.3) \quad \zeta(z) = 2^z \pi^{z-1} \sin\left(\frac{z\pi}{2}\right) \Gamma(1-z) \zeta(1-z)$$

függvényegyenletet. Az (1.3)-ból következik, hogy $\zeta(z)$ nem zérus $Re\ z < 0$ esetén, kivéve a $\{-2m\}_{m=1}^{\infty}$ egyszeres, valós zérushelyeket, a $\zeta(z)$ ún. *triviális zérushelyeit*. Ismeretes továbbá, hogy (lásd [32, 30. oldal]) $\{-2m\}_{m=1}^{\infty}$ jelenti a $\zeta(z)$ összes *valós* gyökeit, és hogy $\zeta(z)$ végtelen sok nem valós gyökkel rendelkezik, amelyek szükségképpen a $0 \leq Re\ z \leq 1$ sávban, vagyis a $\zeta(z)$ ún. *kritikus sávjában* fekszenek. 1859-ben Riemann [30] kimondta híres sejtését, amely a következőképpen hangzik:

$$(1.4) \quad \text{Riemann hipotézis (1859):} \quad \begin{aligned} &A \zeta(z) \text{ valamennyi zérushelye a} \\ &Re\ z = 1/2 \text{ egyenesen fekszik.} \end{aligned}$$

Ezt követően, és ettől függetlenül Hadamard és de la Vallée-Poussin 1896-ban megmutatta (lásd [32, 45. oldal]), hogy $\zeta(z)$ -nek nincs gyöke a $Re\ z = 1$ egyenesen, ami az első bizonyítást szolgáltatta a híres *prímszám tételre*, vagyis a

$$\pi(x) \sim \frac{x}{\log x} \quad (x \rightarrow +\infty)$$

aszimptotikus relációra, ahol $\pi(x)$ jelenti az x -et meg nem haladó prímek számát. Mivel $\zeta(z) \neq 0$ a $\operatorname{Re} z = 1$ egyenesen, (1.3)-ból következik, hogy $\zeta(z)$ a $\operatorname{Re} z = 0$ egyenesen sem tűnik el; következésképpen $\zeta(z)$ -nak végtelen sok nem valós gyöke van a $0 < \operatorname{Re} z < 1$ sávban. Ezekután Hardy [12] 1914-ben megmutatta, hogy $\zeta(z)$ -nek végtelen sok gyöke van a $\operatorname{Re} z = 1/2$ *kritikus egyenesen*, és 1974-ben Levinson [18] bebizonyította, hogy a $\zeta(z)$ kritikus sávbeli gyökeinek legalább egy harmada a kritikus egyenesen fekszik. (1989-ben Conrey [3] az egy harmadot két ötödre javította.)

Az elmúlt évek során, még a szuperszámítógépek feltűnése előtt is, sokan végeztek numerikus számításokat arra vonatkozóan, hogy a kritikus sávban gyököket találjanak. Jegyezzük meg, hogy ha z egy nem valós gyöke a $\zeta(z)$ -nak, akkor (1.1) és (1.3) maga után vonja, hogy \bar{z} , $1 - z$ és $1 - \bar{z}$ is nem valós gyökök. Így a $\zeta(z)$ nem valós gyökeinek keresésénél elegendő a kritikus sáv *felső* félsíkba eső részére szorítkozni, vagyis az

$$(1.5) \quad S := \{z \in \mathbb{C} : 0 < \operatorname{Re} z < 1 \text{ és } \operatorname{Im} z > 0\}$$

tartományra. Van de Lune, te Riele és Winter [19] 1986-ban kiszámították, hogy

$$(1.6) \quad \left\{ \begin{array}{l} \text{a } 0 < \operatorname{Im} z \leq 545439823,215 \text{ egyenlőtlenségekkel definiált } S \\ \text{tartományban pontosan } 1\,500\,000\,001 \text{ gyöke van a } \zeta(z), \\ \text{függvénynek, amelyek mind pontosan a } \operatorname{Re} z = 1/2 \text{ egyenesen} \\ \text{vannak és e gyökök mind egyszeresek.} \end{array} \right.$$

E számítások nagysebességű kompjutereken (elsősorban a Cyber 205-ön) mintegy 1500 óra gépidőt vettek igénybe. Hasonlóan hatékony, mintegy 1000 órát igénylő számításokat végzett Odlyzko [22] a CRAY X-MP szuperszámítógépen, részben a Riemann hipotézis helyességének ellenőrzésére, de elsősorban a $\zeta(z)$ -nek a $\operatorname{Re} z = 1/2$ egyenesen lévő szomszédos gyökei távolságára kimondott sejtésekkel kapcsolatban. E számításokkal előállította a $\zeta(z)$ 78 893 234 szomszédos gyökét az S

$$(1.7) \quad T := \left\{ \begin{array}{l} z \in \mathbb{C} : 0 < \operatorname{Re} z < 1 \text{ és } \alpha \leq \operatorname{Im} z \leq \beta, \text{ ahol} \\ \alpha := 15202440115916180028,24, \text{ és} \\ \beta := 15202440115927890387,66. \end{array} \right.$$

részalmazán. Valamennyi, e T téglalapban talált gyök ismét csak egyszeresnek bizonyult, és mind a $\operatorname{Re} z = 1/2$ kritikus egyenesen volt. Ezekkel a nagyszabású számításokkal mind a mai napig *nem találtak* ellenpéldát a Riemann hipotézisre, s ez bizonyos mértékben azt sugallja, hogy a hipotézis valószínűleg igaz!

Sokrétű *más* matematikai módszerek ismeretese a Riemann hipotézis „megtámadására”. E fejezet soronkövetkező részeiben az a célunk, hogy rámutassunk a zeta-függvény zérushelyeinek segítségével megfogalmazott Riemann hipotézis, az

1927-ben kimondott Pólya-féle sejtés (mostanában kapott megoldásával együtt), valamint a Laguerre- és Turán-féle egyenlőtlenségek szűkebb értelemben vett kapcsolatára.

3.2. A Pólya sejtés

E részben az 1927-ben Pólya György által kimondott sejtéssel foglalkozunk, ami a Riemann hipotézis egy gyengébb formája. Először is definiáljuk a Riemann-féle ξ -függvényt, amely Riemann eredeti jelölésében [30, 147. oldal] (ezt használja Pólya [25] is) a következő:

$$(2.1) \quad \xi(iz) := \frac{1}{2} \left(z^2 - \frac{1}{4} \right) \pi^{-\frac{1}{2} - \frac{1}{4}} \Gamma \left(\frac{z}{2} + \frac{1}{4} \right) \zeta \left(z + \frac{1}{2} \right),$$

ahol ζ az (1.1)-beli Riemann-féle ζ -függvény. (Figyelmeztetjük az olvasót, hogy Titchmarsh [32, 16. oldal] a Ξ szimbólumot használja ξ helyett.) A $\Gamma(z)$ és $\zeta(z)$ ismert gyökeit és pólusait figyelembevéve elmondhatjuk, hogy $\xi(z)$ egy egész függvény, amelyről megmutatható, hogy 1 rendű (lásd [32, 29. oldal]). Ezenkívül a Riemann-féle ξ -függvény előállítható mint egy Fourier-transzformált (lásd [32, 255. oldal]):

$$(2.2) \quad \frac{1}{8} \xi \left(\frac{x}{2} \right) = \int_0^\infty \Phi(t) \cos(xt) dt,$$

ahol, ha

$$(2.3) \quad a_n(t) := (2\pi^2 n^4 e^{9t} - 3\pi n^2 e^{5t}) \exp(-\pi n^2 e^{4t}) \quad (n = 1, 2, \dots),$$

akkor

$$(2.4) \quad \Phi(t) := \sum_{n=1}^{\infty} a_n(t) \quad (t \in \mathbb{R}).$$

Az olvasó kényelmére, valamint az elkövetkezendők kedvéért az alábbiakban kimondjuk Pólya [25] néhány eredményét, amelyek összefoglalják a $\Phi(t)$ függvény több ismert tulajdonságát. A $\Phi(t)$ további tulajdonságai megtalálhatók Csordas, Norfolk és Varga [4, Theorem A] munkájában.

1. TÉTEL (Pólya [25]). A (2.4)-beli $\Phi(t)$ függvényre teljesülnek a következők:

$$(2.5) \quad \left\{ \begin{array}{ll} \text{(i)} & \text{minden egyes } n \geq 1 \text{ esetén } a_n(t) > 0 \text{ minden } t \geq 0 \text{-ra,} \\ & \text{úgyhogy } \Phi(t) > 0 \text{ minden } t \geq 0 \text{ esetén;} \\ \text{(ii)} & \Phi(z) \text{ analitikus } a - \pi/8 < \operatorname{Im} z < \pi/8 \text{ sávban;} \\ \text{(iii)} & \Phi(t) \text{ egy páros függvény, úgyhogy } \Phi^{(2m+1)}(0) = 0 \\ & (m = 0, 1, \dots); \\ \text{(iv)} & \text{bármely } \varepsilon > 0 \text{ esetén } \lim_{t \rightarrow \infty} \Phi^{(n)}(t) \exp[(\pi - \varepsilon)e^{4t}] = 0 \\ & (n = 0, 1, \dots). \end{array} \right.$$

A (2.5i) – (2.5iv) állítások bizonyítása megtalálható Pólya [25] munkájában. A (2.5iii) állítást kivéve, amely egyáltalán nem triviális, a $\Phi(t)$ fenti tulajdonságai közvetlenül beláthatók. Visszatérve a (2.2) alatti integrál-előállításához, sorbafejtve a $\cos(xt)$ függvényt és tagonként integrálva, a $\xi\left(\frac{x}{2}\right)/8$ Maclaurin sorára a következőt kapjuk:

$$(2.6) \quad \frac{1}{8}\xi\left(\frac{x}{2}\right) = \sum_{m=0}^{\infty} \frac{\hat{b}_m(-x^2)^m}{(2m)!},$$

ahol

$$(2.7) \quad \hat{b}_m := \int_0^{\infty} t^{2m} \Phi(t) dt \quad (m = 0, 1, \dots).$$

A (2.6)-ban a $z = -x^2$ helyettesítést alkalmazva, definiáljuk az

$$(2.8) \quad F_0(z) := \sum_{m=0}^{\infty} \frac{\hat{b}_m z^m}{(2m)!}$$

függvényt, úgyhogy (2.6) alapján

$$(2.9) \quad \frac{1}{8} \xi\left(\frac{x}{2}\right) = F_0(-x^2).$$

Mivel ξ egy 1 rendű egész függvény, (2.9) azt mutatja, hogy F_0 egy 1/2 rendű egész függvény. Továbbá, mivel (2.5i) alapján a \hat{b}_m együtthatók mind pozitívak, ezért F_0 egy valós egész függvény (vagyis $F_0(z)$ valós bármely valós z esetén), és $F_0(x) > 0$ bármely $x \geq 0$ mellett. Annak érdekében, hogy a (2.8) alatti F_0 függvényt összekapcsoljuk az (1.4) alatti Riemann hipotézissel, először is érdemes megjegyezni, hogy ez utóbbi eredeti megfogalmazása (lásd Riemann [30, 148. oldal]) szerint a ξ függvény összes gyöke (lásd (2.1)) valós. A (2.9)-ből láthatjuk, hogy x_0 akkor és csak akkor gyöke a $\xi\left(\frac{x}{2}\right)$ függvénynek, ha $z_0 := -x_0^2$ negatív valós gyöke az $F_0(z)$ -nek. Ebből következik, hogy a Riemann hipotézis igazsága ekvivalens azzal az állítással, hogy F_0 valamennyi gyöke valós és negatív.

Mindezekután célszerűnek látszik bevezetni a valós egész függvények következő osztályát. A szokásos $\prod_{j=1}^0 := 1$ jelöléssel élve, tekintsünk egy tetszőleges,

$$(2.10) \quad f(z) = C e^{-\lambda z^2 + \beta z} z^n \prod_{j=1}^{\omega} \left(1 - \frac{z}{x_j}\right) e^{z/x_j} \quad (0 \leq \omega \leq \infty)$$

alakú valós egész függvényt, ahol $\lambda \geq 0$, β és $C \neq 0$ valós konstansok, n egy nemnegatív egész szám, és az x_j -k valós, nullától különböző számok úgy, hogy $\sum_{j=1}^{\omega} x_j^{-2} < \infty$. Az ilyen alakban előállítható függvények összességét *Laguerre-Pólya osztálynak* nevezzük, és az $f \in \mathcal{L} - \mathcal{P}$ jelölést használjuk, amennyiben f a (2.10) alakú. Az alábbi eredmény lényegében Laguerre-től [17, 174. oldal] származik (lásd Boas [1, 24. oldal]).

2. TÉTEL (Laguerre [17]). Legyen $f \in \mathcal{L} - \mathcal{P}$ (lásd (2.10)). Ekkor

$$(2.11) \quad L_{p+1}(x; f) := (f^{(p+1)}(x))^2 - f^{(p)}(x)f^{(p+2)}(x) \geq 0 \quad (x \in \mathbb{R}, p = 0, 1, \dots)$$

Bizonyítás. Az f logaritmikus deriváltja (2.10) alapján

$$\frac{f'(z)}{f(z)} = -2\lambda z + \beta + \frac{n}{z} + \sum_{j=1}^{\omega} \left\{ \frac{1}{(z - x_j)} + \frac{1}{x_j} \right\},$$

amiből az $f \in \mathcal{L} - \mathcal{P}$ feltevést használva következik, hogy

$$(2.12) \quad \left(\frac{f'(x)}{f(x)} \right)' = \frac{f(x)f''(x) - (f'(x))^2}{f^2(x)} = - \left[2\lambda + \frac{n}{x^2} + \sum_{j=1}^{\omega} \frac{1}{(x - x_j)^2} \right] \leq 0,$$

vagyis

$$(2.13) \quad (f'(x))^2 - f(x)f''(x) \geq 0 \quad (x \in \mathbb{R}).$$

Továbbá ismeretes (lásd Pólya és Schur [27]), hogy $f \in \mathcal{L} - \mathcal{P}$ maga után vonja az $f' \in \mathcal{L} - \mathcal{P}$ relációt, amiből (2.11) indukcióval következik. \square

Ha az $f \in \mathcal{L} - \mathcal{P}$ Maclaurin sora $f(z) = \sum_{m=0}^{\infty} c_m z^m$, akkor (2.11)-et $x = 0$ -ra alkalmazva adódik, hogy

$$(2.14) \quad mc_m^2 \geq (m+1)c_{m-1}c_{m+1} \quad (m = 1, 2, \dots).$$

Viszont, ha $\lambda > 0$ vagy $\omega = \infty$ az adott $f \in \mathcal{L} - \mathcal{P}$ esetén, akkor (2.12) miatt (2.13)-ban, és így (2.14)-ben is, a szigorú egyenlőtlenség érvényesül, vagyis

$$(2.15) \quad mc_m^2 > (m+1)c_{m-1}c_{m+1} \quad (m = 1, 2, \dots).$$

Ezt a következőképpen használhatjuk fel.

Mivel F_0 egy valós, $1/2$ -rendű egész függvény, a Riemann hipotézis maga után vonná, hogy az F_0 végtelen sok gyöke mind valós és negatív volna. Emiatt a (2.8) alatti F_0 -ra alkalmazva az erősebb (2.15) egyenlőtlenségeket, kapnánk, hogy

$$(2.16) \quad (\hat{b}_m)^2 > \left(\frac{2m-1}{2m+1} \right) \hat{b}_{m-1}\hat{b}_{m+1} \quad (m = 1, 2, \dots),$$

vagyis

$$(2.17) \quad D_m := (\hat{b}_m)^2 - \left(\frac{2m-1}{2m+1} \right) \hat{b}_{m-1}\hat{b}_{m+1} > 0 \quad (m = 1, 2, \dots).$$

Ílymódon láthatjuk, hogy *valamennyi* (2.16) alatti egyenlőtlenség *szükséges* feltétele a Riemann hipotézis igazságának. Modern terminológiát alkalmazva, a (2.16) egyenlőtlenségeket általában *Turán-féle egyenlőtlenségeknek*, míg a (2.17) alatti D_m számokat *Turán-féle differenciáknak* nevezik.

1925–1926-ban, röviddel Jensen 1925-ben bekövetkezett halála után Pólyának lehetősége nyílt arra, hogy tanulmányozza Jensen kiadatlan jegyzeteit. A matematikai társadalom szempontjából ez nagy jelentőséggel bírt, hiszen Jensen 14 évvel korábban bejelentette, hogy új, a Riemann hipotézisre vonatkozó szükséges és elégséges feltételeket fog publikálni, de ez soha nem következett be. 1927-ben Pólya [25] publikálta Jensen jegyzeteit, és ez tartalmazta a fenti szükséges és elégséges feltételeket. Cikkében Pólya használta a

$$b_m := \frac{2 \cdot m!}{(2m)!} \hat{b}_m \quad (m = 0, 1, \dots)$$

jelölést, ahol \hat{b}_m a (2.7) által definiált, és így e jelöléssel az F_0 -ra alkalmazott (2.14) gyengébb formában a következő lesz:

$$b_n^2 - b_{n-1}b_{n+1} \geq 0 \quad (n = 1, 2, \dots).$$

Pólya [25, 16. oldal] a következő megjegyzést is tette:¹

„Mindazonáltal, hogy mily keveset haladtunk ebben az irányban, azt világosan mutatja [az a tény], hogy még a ξ -függvény gyökeinek valós voltára vonatkozó $b_n^2 - b_{n-1}b_{n+1} \geq 0$, majdhogynem legegyszerűbb szükséges feltétel sem nyert mindezeig bizonyítást.”

E megjegyzés az alábbi formában vált közismertté:

(2.18) A Pólya sejtés (1927):

$$(\hat{b}_m)^2 \stackrel{?}{>} \left(\frac{2m-1}{2m+1} \right) \hat{b}_{m-1} \hat{b}_{m+1} \quad (m = 1, 2, \dots).$$

A Pólya-sejtés iránti érdeklődés teljesen természetes: ha a (2.18) alatti egyenlőtlenségek valamelyike *nem teljesülne* egy $m \geq 1$ mellett, vagy ekvivalens módon, ha a (2.17)-beli D_m *nem lenne pozitív* valamely $m \geq 1$ esetén, akkor a Riemann hipotézis *hamisnak bizonyulna*.

Érdeemes felidézni az 1927-ben kimondott Pólya-sejtés történetét. Szemmel láthatóan ez a probléma érintetlen maradt közel 40 évig. Majd 1966 és 1969-ben

¹ „Wie wenig wir aber an diesem Wege fortgeschritten sind, zeigt es am deutlichsten, dass die nächst einfachste notwendige Bedingung für die Realität aller Nullstellen

$$(17) \quad b_n^2 - b_{n-1}b_{n+1} \geq 0$$

im Falle der ξ -Funktion noch nicht verifiziert ist.”

Grosswald [10,11] általánosította Hayman [13] egy, a megengedett függvényekre vonatkozó formuláját, és ennek alkalmazásaképpen, a (2.16) és (2.17) jelöléssel élve bebizonyította, hogy

$$(2.19) \quad \begin{aligned} D_m &:= \left(\hat{b}_m\right)^2 - \left(\frac{2m-1}{2m+1}\right) \hat{b}_{m-1} \hat{b}_{m+1} \\ &= \frac{\left(\hat{b}_m\right)^2}{m} \left\{ 1 + O\left(\frac{1}{\log m}\right) \right\} \quad (m \rightarrow \infty). \end{aligned}$$

Mivel a \hat{b}_m momentumok (2.5i) miatt szükségsszerűen pozitívok minden $m \geq 0$ esetén, Grosswald (2.19) alatti eredménye azt bizonyítja, hogy (2.16) igaz minden elég nagy m -re, mondjuk $m \geq m_0$ mellett. Sajnos, Grosswald érvelése nem adja meg m_0 értékét, és – tudomásunk szerint – ezt a hiányt a Pólya-sejtés megoldásában azóta sem pótolták.

A (2.16) alatti Turán-egyenlőtlenségek mindent eldöntő természete a következő megfontolásból is látható. Mivel (2.5i) alapján $\Phi(t)$ pozitív minden $t \geq 0$ mellett, a (2.7)-beli \hat{b}_m momentumot írhatjuk a

$$\hat{b}_m^2 = \left\{ \int_0^\infty t^{(2m-2)/2} \sqrt{\Phi(t)} \cdot t^{(2m+2)/2} \sqrt{\Phi(t)} dt \right\}^2$$

alakban. Az itt szereplő integrálra alkalmazva a Cauchy-Schwarz egyenlőtlenséget, közvetlenül kapjuk, hogy $\left(\hat{b}_m\right)^2 \leq \hat{b}_{m-1} \hat{b}_{m+1}$, ami ekvivalens a

$$(2.20) \quad \hat{b}_m^2 \leq \left(\frac{2m+1}{2m-1}\right) \hat{b}_{m-1} \hat{b}_{m+1} \quad (m = 1, 2, \dots),$$

egyenlőtlenségekkel. Így a keresett (2.16) alatti Turán-egyenlőtlenségek ennek majd-hogynem fordítottjai:

$$\hat{b}_m^2 > \left(\frac{2m-1}{2m+1}\right) \hat{b}_{m-1} \hat{b}_{m+1} \quad (m = 1, 2, \dots).$$

A Pólya-sejtést mi (mármint G. Csordas, T. S. Norfolk és jómagam) főképp azért tartottuk csodálatosnak, mert a (2.16) alatti Turán-egyenlőtlenségeket numerikusan senki sem ellenőrizte. Márpedig a (2.7) alatti \hat{b}_m momentumok nagypontosságú kiszámítása nem látszik lehetetlennek, hiszen (2.5iv) szerint $\Phi(t)$ rendkívül gyorsan tart nullához, midőn $t \rightarrow \infty$. Ezért 1983-ban lelkesen megkezdtük a $\left\{\hat{b}_m\right\}_{m=0}^{109}$ momentumok nagypontosságú (50 értékes jegy) kiszámítását, amiből azután meghatároztuk a Turán-féle $\{D_m\}_{m=1}^{108}$ differenciákat. Végül is volt *valami* esély arra, hogy egy $D_m \leq 0$ egyenlőtlenségre bukkanunk.

Sajnos, számításaink nem produkáltak meglepő eredményt: az összes kiszámított D_m pozitívnak bizonyult. A 3.1. Táblázatban felsoroljuk a $\{\hat{b}_m\}_{m=0}^{20}$, $\{D_m\}_{m=1}^{20}$ és $\{\tilde{D}_m\}_{m=1}^{20}$ számokat, ahol

$$(2.21) \quad \tilde{D}_m := \frac{mD_m}{(\hat{b}_m)^2} \quad (m = 1, 2, \dots),$$

úgyhogy (lásd (2.19))

$$(2.21') \quad \tilde{D}_m = 1 + O\left(\frac{1}{\log m}\right) \quad (m \rightarrow \infty).$$

A 3.1. Táblázat összes számait 11 értékes jegyre kerekítettük. Mivel a \hat{b}_m momentumok mind pozitívak, a (2.21)-ből láthatjuk, hogy \tilde{D}_m pozitív akkor és csak akkor, ha a D_m Turán-differencia pozitív. Jegyezzük meg, hogy a 3.1. Táblázatban a \tilde{D}_m -ek mint m függvénye sokkal lassabban változnak, mint a megfelelő D_m értékek, ami konzisztens a (2.21') alatti eredménnyel.

Úgy tűnhet a 3.1. Táblázatból, hogy a \hat{b}_m momentumok szigorúan monoton csökkennek, ha $m \geq 0$. Könnyen látható, hogy ez nincs így, és ebből következőleg egy kis matematikai probléma adódott: határozzuk meg, numerikusan és elméletileg (aszimptotikus módszereket alkalmazva), azt a pontos m értéket, amelyre \hat{b}_m minimális. Kiderült (lásd Csordas, Norfolk, és Varga [5]), hogy \hat{b}_m szigorúan monoton csökkenő, ha $0 \leq m \leq 339$, és szigorúan monoton növekedő, ha $m \geq 339$; úgyhogy

$$(2.22) \quad \min_{m \geq 0} \hat{b}_m = \hat{b}_{339} \doteq 2,18540\,10467 \dots 10^{-71}.$$

E rész befejezéseként még két megjegyzést teszünk. Először is, a $\{\hat{b}_m\}_{m=0}^{109}$ momentumok kiszámításának részleteit a 3.6 részben fogjuk tárgyalni. Másodszor, a 3.1. Táblázatbeli számok numerikus meghatározása nem volt hiábavaló. Mint azt a 3.3 részben látni fogjuk, ezek a numerikus számítások motiváltak egy pontos bizonyítást a Pólya-sejtésre, és még más dolgokat is!

3.3. A Pólya-sejtés egy újabb megoldása

1986-ban Csordas, Norfolk és Varga [4] egy hosszú (12 lemmából álló) bizonyítást adott arra a tényre (lásd (2.4)), hogy ha

$$(3.1) \quad \Psi(t) := \int_t^\infty \Phi(\sqrt{u}) \, du \quad (t \geq 0),$$

akkor

$$(3.2) \quad \log \Psi(t) \text{ szigorúan konkáv a } (0, +\infty) \text{ intervallumon.}$$

Ezzel [4]-ben konstruktív módon bebizonyítottuk a közel hatvan éves Pólya sejtést! Céljainknak megfelelően, az alábbiakban ismertetjük Csordas és Varga [7] egy további, a (3.2)-höz hasonló eredményét, amely azonban alapjául szolgál egy, a Pólya sejtésnél erősebb állításnak. Kiindulópontunk az alábbi tétel lesz.

m	\hat{b}_m	D_m	\tilde{D}_m
0	6,21400 97273 (-2)	—	—
1	7,17873 25984 (-4)	3,58844 91486 (-8)	6,96323 80609 (-2)
2	2,31472 53388 (-5)	3,16329 93950 (-11)	1,18078 64542 (-1)
3	1,17049 98956 (-6)	7,05673 24419 (-14)	1,54519 91985 (-1)
4	7,85969 60229 (-8)	2,83222 02230 (-16)	1,83389 94114 (-1)
5	6,47444 26609 (-9)	1,73636 66894 (-18)	2,07112 67219 (-1)
6	6,24850 92806 (-10)	1,47803 17201 (-20)	2,27134 00906 (-1)
7	6,85711 35660 (-11)	1,64153 36845 (-22)	2,44379 71530 (-1)
8	8,37956 28564 (-12)	2,27744 38477 (-24)	2,59474 65240 (-1)
9	1,12289 59005 (-12)	3,82273 77260 (-26)	2,72858 83427 (-1)
10	1,63076 65724 (-13)	7,57537 75877 (-28)	2,84852 92500 (-1)
11	2,54307 50583 (-14)	1,73849 34268 (-29)	2,95697 31508 (-1)
12	4,22669 38654 (-15)	4,54925 56467 (-31)	3,05576 53485 (-1)
13	7,44135 71845 (-16)	1,34019 54348 (-32)	3,14635 11038 (-1)
14	1,38066 04233 (-16)	4,39776 86757 (-34)	3,22988 20677 (-1)
15	2,68793 65964 (-17)	1,59301 19382 (-35)	3,30728 97809 (-1)
16	5,47056 43869 (-18)	6,32085 57309 (-37)	3,37933 76956 (-1)
17	1,16018 31858 (-18)	2,72899 35268 (-38)	3,44665 87993 (-1)
18	2,55669 85949 (-19)	1,27457 93250 (-39)	3,50978 33473 (-1)
19	5,84001 96623 (-20)	6,40679 74312 (-41)	3,56915 96545 (-1)
20	1,37967 28720 (-20)	3,45025 04583 (-42)	3,62516 99207 (-1)

9.1. Táblázat

3. TÉTEL ([7]). A (2.4)-beli $\Phi(t)$ függvény rendelkezik azzal a tulajdonsággal, hogy

$$(3.3) \quad \log \Phi(\sqrt{t}) \text{ szigorúan konkáv a } (0, +\infty) \text{ intervallumon.}$$

Bizonyítás (vázlat). Az 1. Tétel (2.5ii) része szerint $\Phi(z)$ analitikus a $-\pi/8 < \operatorname{Im} z < \pi/8$ sávban, ezért $\Phi(t) \in C^\infty(\mathbb{R})$, és mivel (2.5i) alapján $\Phi(t) > 0$ bármely $t \geq 0$ mellett, innen könnyen látható, hogy $(d^2/dt^2) \log \Phi(\sqrt{t}) < 0$ a $(0, +\infty)$ intervallumon akkor és csak akkor, ha $g(t) > 0$ a $(0, +\infty)$ intervallumon. Ebből következik, hogy

$$(3.4) \quad g(t) := t \left[(\Phi'(t))^2 - \Phi(t)\Phi''(t) \right] + \Phi(t)\Phi'(t) \quad (t \geq 0).$$

A [7]-ben egy hosszadalmas, tíz lemmából álló konstrukcióval (ezt itt nem ismertetjük) korlátokat találtunk a $\Phi^{(j)}(t)$ ($j = 1, 2, \dots, 6$) függvényekre különböző intervallumokon, s ezek alapján

$$(3.5) \quad \begin{cases} g(t) > 0 & (0 < t \leq 0.03), \\ g(t) > 0 & (0.03 \leq t \leq 0.06), \\ g(t) > 0 & (0.056 \leq t < \infty), \end{cases}$$

ami a kívánt (3.3) eredményt adja. \square

Ezen a helyen megemlítjük, hogy korábban, 1982-ben (erről akkor nem tudtunk) Matiyasevich [20] egy hasznos, hármas integrált tartalmazó formulát adott a (2.17) alatti D_m Turán-féle differenciákra, nevezetesen

$$(3.6) \quad 2(2m+1)D_m = \int_0^\infty \int_0^\infty u^{2m} v^{2m} \Phi(u) \Phi(v) \left[(v^2 - u^2) \int_u^v -\frac{d}{dt} \left(\frac{\Phi'(t)}{t\Phi(t)} \right) dt \right] du dv$$

bármely $m \geq 1$ esetén, ami a $g(t)$ (3.4)-beli definíciója alapján úgy is írható, hogy

$$(3.7) \quad 2(2m+1)D_m = \int_0^\infty \int_0^\infty u^{2m} v^{2m} \Phi(u) \Phi(v) \left[(v^2 - u^2) \int_u^v \frac{g(t)dt}{t^2 \Phi^2(t)} \right] du dv,$$

bármely $m \geq 1$ esetén. Mivel $g(t)$ és $\Phi(t)$ pozitívak a $(0, +\infty)$ intervallumon (lásd (3.5) és (2.5i)), világos, hogy (3.7)-ben a zárójelbeli kifejezés nemnegatív bármely $0 \leq u, v < \infty$ esetén. De mivel a (3.7) alatti integrandusban a többi tényező is nemnegatív, így $D_m > 0$ bármely $m \geq 1$ esetén. Tehát (lásd (2.12) és (2.13)) a *Pólya sejtés igaz*. Érdekes megemlíteni, hogy Matiyasevich [20] szerint a (3.4)-beli $g(t)$ pozitív a $(0, +\infty)$ intervallumon, „intervallum-számítások szerint, amelyek ugyanolyan meggyőzőek, mint egy bizonyítás”, és, míg a $g(t) > 0$ egyenlőtlenséget demonstráló intervallum-aritmetikai számítások hasonlóképpen részekre osztják a $(0, +\infty)$ intervallumot, mint (3.5)-ben, képtelenek voltak megfelelő információt kapni [20]-ból ahhoz, hogy teljesen meggyőződjünk a szerző számításainak helyességéről. Ennek ellenére, a $g(t) > 0$ egyenlőtlenség a $(0, +\infty)$ intervallumon természetesen igaz, hiszen ennek konstruktív bizonyítása megtalálható [7]-ben, továbbá a (3.1)-beli $\Psi(t)$ logaritmikus konkávitása bizonyított [4]-ben. Hangsúlyozzuk, hogy a Pólya sejtés általunk adott újabb bizonyítása azon a független tényen alapult, hogy $\log \Phi(\sqrt{t})$ szigorúan konkáv a $(0, +\infty)$ intervallumon, és ez utóbbit használni fogjuk az általánosabb eredmény, nevezetesen az e részben tárgyalandó 4. Tétel bizonyításánál.

Mármost tegyük fel, hogy $K(t) : \mathbb{R} \rightarrow \mathbb{R}$ kielégíti a következő feltételeket:

$$(3.8) \quad \begin{cases} (i) & K \text{ integrálható } \mathbb{R} \text{-en,} \\ (ii) & K(t) > 0 \quad (t \in \mathbb{R}), \\ (iii) & K(t) = K(-t) \quad (t \in \mathbb{R}), \text{ és} \\ (iv) & \text{valamely } \varepsilon > 0 \text{ esetén } K(t) = O(\exp(-|t|^{2+\varepsilon})) \quad (t \rightarrow \infty). \end{cases}$$

Az ilyen $K(t)$ függvényeket *megengedett magfüggvényeknek* nevezzük. Ismeretes (lásd Pólya [26]), hogy egy megengedett $K(t)$ magfüggvény

$$(3.9) \quad H(x, K) := \int_{-\infty}^{+\infty} K(t) e^{ixt} dt = 2 \int_0^\infty K(t) \cos(xt) dt$$

Fourier transzformáltja egy valós, véges ρ rendű egész függvény, ahol a (3.8 iv)-beli ε -nal ez a ρ kielégíti a

$$(3.10) \quad \rho \leq \frac{2 + \varepsilon}{1 + \varepsilon} < 2$$

feltételt. Továbbá, ha $K(t)$ egy megengedett magfüggvény és a (3.9)-beli $H(x, K)$ csak *valós* gyökökkel rendelkezik, akkor Pólya [26] egy másik szép eredménye szerint a

$$(3.11) \quad J(x; K, f) := \int_{-\infty}^{+\infty} f(it)K(t)e^{ixt} dt \quad (f \in \mathcal{L} - \mathcal{P})$$

egész függvénynek *szintén csak valós gyökei vannak*. Másszóval, Pólya terminológiájával élve (lásd [26, 7. oldal]), az $f(it)$ függvények *univerzális tényezők*, amelyek megtartják a (3.9)-cel definiált $H(x, K)$ egész függvény gyökeinek valós voltát. (Valójában Pólya [26] megmutatta, hogy az $f(it)$ függvények, ahol $f \in \mathcal{L} - \mathcal{P}$, az egyetlen *analitikus* függvények, amelyek bírnak e tulajdonsággal!)

A fentieket a következőképpen alkalmazhatjuk. Először is nyilvánvaló (2.5)-ből, hogy $\Phi(t)$ kielégíti (3.8)-at, úgyhogy $\Phi(t)$ egy megengedett magfüggvény. Ezért a (3.9) jelöléssel élve kapjuk, hogy (lásd (2.2))

$$(3.12) \quad H(x, \Phi) = 2 \int_0^{\infty} \Phi(t) \cos(xt) dt = \frac{1}{4} \xi\left(\frac{x}{2}\right).$$

Így a Riemann hipotézis helyessége maga után vonná, hogy $H(x, \Phi)$ csak valós gyökökkel rendelkezik. Mármint tekintsünk bármely *páros* $f \in \mathcal{L} - \mathcal{P}$ függvényt, vagyis (2.10) alapján legyen

$$(3.13) \quad f(z) = C e^{-\lambda z^2} z^{2n} \prod_{j=1}^{\omega} \left(1 - \frac{z^2}{x_j^2}\right) \quad (0 \leq \omega \leq \infty),$$

ahol $\lambda \geq 0$, C egy zérustól különböző valós konstans, n egy nemnegatív egész, és az x_j -k kielégítik az $x_j > 0$ és $\sum_{j=1}^{\omega} 1/x_j^2 < \infty$ egyenlőtlenségeket. De egy előjelváltozástól eltekintve könnyen megmutatható, hogy $f(it)\Phi(t)$ egy megengedett magfüggvény *bármely* páros $f \in \mathcal{L} - \mathcal{P}$ esetén (lásd (3.13)). (A bizonyításhoz a (2.5) tulajdonságai használhatók fel.) Következésképpen

$$(3.14) \quad J(x; \Phi, f) := 2 \int_0^{\infty} f(it) \Phi(t) \cos(xt) dt$$

egy 2-nél kisebb rendű valós egész függvény (lásd (3.10)). Továbbá a fenti megfontolásokból következik, hogy a Riemann hipotézis helyessége maga után vonná,

hogy $J(x; \Phi, f)$ csak valós gyökökkel rendelkezik bármely páros $f \in \mathcal{L} - \mathcal{P}$ esetén. Mármint a $J(x; \Phi, f)$ Maclaurin kifejtése a következő:

$$(3.15) \quad J(x; \Phi, f) = \sum_{m=0}^{\infty} \frac{\hat{b}_m(f)(-x^2)^m}{(2m)!},$$

ahol

$$(3.16) \quad \hat{b}_m(f) := 2 \int_0^{\infty} t^{2m} f(it) \Phi(t) dt \quad (m = 0, 1, \dots),$$

és (3.15)-ben elvégezve a $z = -x^2$ helyettesítést, a (2.8) analógiájára kapjuk az

$$(3.17) \quad F(z; f) := \sum_{m=0}^{\infty} \frac{\hat{b}_m(f)z^m}{(2m)!}$$

valós egész függvényt, amelynek rendje (3.10) alapján kisebb, mint 1. Viszont ekkor a 2. Tételt alkalmazva a $F(z; f)$ függvényre, a Riemann hipotézisből következne, hogy

$$(3.18) \quad D_m(f) := \left(\hat{b}_m(f) \right)^2 - \left(\frac{2m-1}{2m+1} \right) \hat{b}_{m-1}(f) \hat{b}_{m+1}(f) > 0 \quad (m = 1, 2, \dots)$$

bármely ilyen páros $f(z) \in \mathcal{L} - \mathcal{P}$ függvényre. Megjegyezzük, hogy (3.18) egy általánosított formája a (2.18) Pólya-sejtésnek, mivel az $f(z) \equiv 1$ választás (3.18)-ban (2.18)-at eredményezi.

A 3. Tételben szereplő $\Phi(\sqrt{t})$ függvény (lásd (3.3)) logaritmikus konkávitása alapján kimondhatjuk Csordas és Varga [7] alábbi eredményét.

4. TÉTEL ([7]). Bármely páros, valós,

$$(3.19) \quad h(z) := C e^{-\lambda z^2} z^{2n} \prod_{j=1}^{\omega} \left(1 - \frac{z^2}{x_j^2} \right) \quad (0 \leq \omega \leq \infty)$$

alakú egész függvényre, ahol λ és $C \neq 0$ valós konstansok, n egy nemnegatív egész, és az $x_j - k$ valósak és zérustól különbözőek úgy, hogy $\sum_{j=1}^{\omega} 1/x_j^2 < \infty$, legyen

$$(3.20) \quad \hat{b}_m(h) := 2 \int_0^{\infty} t^{2m} h(it) \Phi(t) dt \quad (m = 0, 1, \dots).$$

Ekkor a megfelelő,

$$(3.21) \quad D_m(h) := \left(\hat{b}_m(h) \right)^2 - \left(\frac{2m-1}{2m+1} \right) \hat{b}_{m-1}(h) \hat{b}_{m+1}(h) \quad (m = 1, 2, \dots)$$

által definiált Turán-féle különbségek kielégítik a

$$(3.22) \quad D_m(h) > 0 \quad (m = 1, 2, \dots)$$

egyenlőtlenséget.

Bizonyítás. A feltevés szerint a

$$(3.23) \quad K(t) := h(it)\Phi(t) \quad (t \in \mathbb{R})$$

(előjeltől eltekintve) egy megengedett magfüggvény, s így (3.7) és (3.3) alapján elegendő megmutatni, hogy

$$(3.24) \quad \log K(\sqrt{t}) \text{ szigorúan konkáv a } (0, +\infty) \text{ intervallumon.}$$

Nyilván

$$K(\sqrt{t}) = \tilde{C} e^{\lambda t} t^n \prod_{j=1}^{\omega} \left(1 + \frac{t}{x_j^2}\right) \Phi(\sqrt{t}) \quad (\tilde{C} := (-1)^n C),$$

úgyhogy

$$\log K(\sqrt{t}) = \log \tilde{C} + \lambda t + n \log t + \sum_{j=1}^{\omega} \log \left(1 + \frac{t}{x_j^2}\right) + \log \Phi(\sqrt{t}).$$

Differenciálással kapjuk, hogy

$$\frac{d^2}{dt^2} \log K(\sqrt{t}) = - \left[\frac{n}{t^2} + \sum_{j=1}^{\omega} \frac{1}{(x_j^2 + t)^2} \right] + \frac{d^2}{dt^2} \log \Phi(\sqrt{t}) \quad (t > 0).$$

Mivel itt a szögletes zárójelbeli kifejezés nemnegatív bármely $t > 0$ esetén, és mivel az utolsó tag (3.3) alapján negatív bármely $t > 0$ esetén, ezért (3.24)-et bebizonyítottuk. \square

Fontos megjegyezni, hogy a 4. Tétel megenged bármely valós λ értéket, míg a fenti, páros univerzális faktorokkal kapcsolatos okoskodás csak $\lambda \geq 0$ mellett igaz.

A 4. Tétel eredményei alkalmazhatók meglehetősen általános feltételek mellett. Valóban, tekintsünk egy tetszőleges,

$$(3.25) \quad \hat{H}(x) := \int_0^{\infty} K(t) \cos(xt) dt$$

alakú egész függvényt, ahol $K(t)$ bármely $C^2(\mathbb{R})$ függvény, amely egy megengedett magfüggvény (lásd (3.8)). Legyenek a $K(t)$ függvénynek megfelelő momentumok

$$(3.26) \quad \hat{c}_m := \int_0^{\infty} t^{2m} K(t) dt \quad (m = 0, 1, 2, \dots).$$

Ekkor annak *szükséges feltétele*, hogy a $\hat{H}(x)$ egész függvénynek csak valós gyökei legyenek az, hogy

$$(3.27) \quad \hat{c}_m^2 > \left(\frac{2m-1}{2m+1} \right) \hat{c}_{m-1} \hat{c}_{m+1} \quad (m = 1, 2, \dots).$$

A 4. Tétel alapján (3.27) fennállásának egy *elégséges feltétele*

$$(3.28) \quad \frac{d^2}{dt^2} \log \left(K(\sqrt{t}) \right) < 0 \quad (t > 0).$$

A (3.28) alkalmazására példaként tekintsük, csakúgy, mint [7]-ben, a $\hat{K}(t) := \exp(-2 \cosh(t))$ függvényt. Ismeretes (lásd Pólya [24]), hogy a $\hat{K}(t)$ koszinusz-transzformáltja, vagyis

$$\int_0^\infty \exp(-2 \cosh(t)) \cos(xt) dt,$$

egy valós egész függvény, amelynek csak valós gyökei vannak. Mivel

$$\log \hat{K}(\sqrt{t}) = -2 \cosh(\sqrt{t}) \quad (t \geq 0),$$

azért

$$\frac{d^2}{dt^2} \log \hat{K}(\sqrt{t}) = \frac{1}{2} \left\{ \frac{\sinh(\sqrt{t})}{t^{3/2}} - \frac{\cosh(\sqrt{t})}{t} \right\} \quad (t > 0).$$

Másrészt, mivel itt a zárójelben lévő kifejezés $u := \sqrt{t}$ változó szerinti Maclaurin sorának együtthatói mind *negatívak*, láthatjuk, hogy a (3.27) egyenlőtlenségekre vonatkozó (3.28) szükséges feltételek teljesülnek. Érdemes megjegyezni, hogy a $\hat{K}(t) = \exp(-2 \cosh(t))$ magfüggvény nem fejezhető ki $f(it)$ alakban, ahol $f(z)$ (2.10) alakú, úgyhogy ebben a példában még univerzális faktorok sem szerepelnek.

A fenti eredmények alkalmazásaképpen megfogalmazzuk az alábbi korolláriumot.

5. KOROLLÁRIUM ([7]). *Legyen*

$$(3.29) \quad K_\lambda(t) := \Phi(t) \cosh(\lambda t) \quad (\lambda \in \mathbb{R}),$$

ahol $\Phi(t)$ a (2.4) által definiált, és legyen

$$(3.30) \quad \hat{c}_m(\lambda) := \int_0^\infty t^{2m} K_\lambda(t) dt \quad (m = 0, 1, \dots).$$

Ekkor

$$(3.31) \quad (\hat{c}_m(\lambda))^2 > \left(\frac{2m-1}{2m+1} \right) \hat{c}_{m-1}(\lambda) \hat{c}_{m+1}(\lambda) \quad (m = 1, 2, \dots; \lambda \in \mathbb{R}).$$

Bizonyítás. A (3.31) egyenlőtlenségek igazolásához elegendő megmutatni, hogy

$$\frac{d^2}{dt^2} \log K_\lambda(\sqrt{t}) = \frac{d^2}{dt^2} \log \Phi(\sqrt{t}) + \frac{d^2}{dt^2} \log(\cosh(\lambda\sqrt{t})) < 0 \quad (0 < t < \infty).$$

Viszont (3.3) szerint ehhez elég $\frac{d^2}{dt^2} \log(\cosh(\lambda\sqrt{t})) \leq 0$ fennállása minden $0 < t < \infty$ és minden $\lambda \in \mathbb{R}$ esetén. Mint az könnyen kiszámítható, ez akkor és csak akkor igaz, ha (az $u := \sqrt{t}$ helyettesítéssel)

$$(3.32) \quad \sigma_\lambda(u) := -\lambda^2 u + \lambda \sinh(2\lambda u)/2 \geq 0 \quad (0 < u < \infty; \lambda \in \mathbb{R}).$$

Azonban $\sigma_\lambda(0) = 0$ és $\sigma'_\lambda(u) = \lambda^2 \{-1 + \cosh(2\lambda u)\} \geq 0$, azért (3.32) fennáll. \square

A (3.31) egyenlőtlenségek ismertek a $\lambda \geq 1$ (lásd Pólya [24, 32. oldal]), valamint $\lambda = 0$ (lásd [4]) speciális esetekben. A $\lambda = 1$ esetén a (3.29)-beli $K_1(t)$ magfüggvény különösen érdekes, mivel Pólya [25] megmutatta, hogy a $K_1(t)$ Fourier koszinusz transzformáltja, vagyis

$$(3.33) \quad F_1(x) := \int_0^\infty K_1(t) \cos(xt) dt,$$

csakis valós gyökökkel rendelkezik. Pólya bizonyítási módszere azt is mutatja, hogy az

$$F_\lambda(x) := \int_0^\infty K_\lambda(t) \cos(xt) dt$$

egész függvénynek csak valós gyökei vannak, ha $\lambda \geq 1$, és ebből következőleg (3.31) teljesül, ha $\lambda \geq 1$.

3.4. A de Bruijn–Newman-féle Λ konstans

Emlékeztetünk arra, hogy a Riemann-féle ξ -függvény Fourier transzformáltja (lásd (2.2)) előállítható az

$$(4.1) \quad \frac{1}{8} \xi\left(\frac{x}{2}\right) = \frac{1}{2} \int_{-\infty}^{\infty} \Phi(t) e^{ixt} dt = \int_0^\infty \Phi(t) \cos(xt) dt,$$

alakban, ahol

$$(4.2) \quad \Phi(t) := \sum_{n=1}^{\infty} (2\pi^2 n^4 e^{9t} - 3\pi n^2 e^{5t}) \exp(-\pi n^2 e^{4t}).$$

Bármely valós λ esetén beiktatjuk az $e^{\lambda t^2}$ tényezőt a (4.1) alatti integrandusba, és legyen

$$(4.3) \quad H_\lambda(x) := \frac{1}{2} \int_{-\infty}^{+\infty} e^{\lambda t^2} \Phi(t) e^{ixt} dt = \int_0^\infty e^{\lambda t^2} \Phi(t) \cos(xt) dt,$$

úgyhogy (2.2) alapján $H_0(x) = \xi(x/2)/8$. Sorbafejtve a $\cos(xt)$ függvényt, a $H_\lambda(x)$ Maclaurin sorát megadhatjuk a

$$(4.4) \quad H_\lambda(x) = \sum_{m=0}^{\infty} \frac{\hat{b}_m(\lambda)(-x^2)^m}{(2m)!}$$

alakban, ahol

$$(4.5) \quad \hat{b}_m(\lambda) := \int_0^{\infty} t^{2m} e^{\lambda t^2} \Phi(t) dt \quad (m = 0, 1, \dots; \lambda \in \mathbb{R}).$$

Ismeretes (lásd Csordas, Norfolk és Varga [6, Appendix A]), hogy $H_\lambda(x)$ egy valós, 1 rendű (és ∞ típusú) egész függvény *bármely* valós λ esetén. Továbbá, mivel $\Phi(t) > 0$ bármely valós t mellett (lásd (2.5i)), ezért $\hat{b}_m(\lambda) > 0$ minden $m = 0, 1, \dots$ és valós λ esetén. Következésképpen a (4.4)-beli $H_\lambda(x)$ egy *valós*, 1 rendű egész függvény bármely valós λ esetén.

E részben az a célunk, hogy a $H_\lambda(x)$ gyökeinek *viselkedését* vizsgáljuk, mint a valós λ paraméter függvényét, és feltárjuk ennek kapcsolatát a Riemann hipotézissel. Először is megjegyezzük, hogy $g(t) := e^{-\lambda t^2}$, $\lambda \geq 0$, eleme az $\mathcal{L} - \mathcal{P}$ -nek (lásd (2.10)), úgyhogy

$$g(it) = e^{\lambda t^2} \quad (\lambda \geq 0)$$

egy *univerzális faktor* Pólya terminológiájában. Ezért Pólya [26] szerint, ha $H_0(x) = \xi(x/2)/8$ csak valós gyökökkel rendelkezik, akkor ugyanez igaz a (4.3)-beli $H_\lambda(x)$ -re is, bármely $\lambda \geq 0$ esetén. Ebből következőleg de Bruijn [2] két eredménye szerint

- (i) $H_\lambda(x)$ csak valós gyökökkel rendelkezik, hacsak $\lambda \geq 1/2$;
 (4.6) (ii) ha $H_\lambda(x)$ csak valós gyökökkel bír valamely valós λ mellett,
 akkor $H_{\lambda'}(x)$ csak valós gyökökkel bír bármely $\lambda' \geq \lambda$ esetén.

Másszóval, a Riemann hipotézis maga után vonná azt, hogy $H_\lambda(x)$ csak valós gyökökkel bír bármely $\lambda \geq 0$ esetén.

Érdekes, hogy Newman [21] nemrég megmutatta, hogy létezik egy valós Λ szám, amelyre $-\infty < \Lambda \leq 1/2$, és

$$(4.7) \quad \begin{aligned} &H_\lambda(x) \text{ CSAK VALÓS GYÖKÖKKEL RENDELKEZIK,} \\ &\text{HA } \lambda \geq \Lambda, \text{ ÉS } H_\lambda(x) \text{ RENDELKEZIK NEM VALÓS} \\ &\text{GYÖKÖKKEL, HA } \lambda < \Lambda. \end{aligned}$$

N. G. de Bruijn idevágó korábbi munkái alapján jogos ezt a Λ konstans *de Bruijn-Newman konstansnak* nevezni.

Newman [21] bizonyítási módszere a Λ konstans *létezésére* nem konstruktív, s ezért [21]-ben nincs explicit *alsó korlát* a Λ -ra. Ami a Λ *felső korlátait* illeti,

$\Lambda \leq 1/2$ bizonyára igaz, és a Riemann hipotézisből következne, hogy $\Lambda \leq 0$, azaz a megjegyzéssel, hogy „ez az új sejtés annak a kijelentésnek a kvantitatív formája, hogy ha a Riemann hipotézis igaz, akkor alig az”. Így elérkeztünk ahhoz az új erőpróbát jelentő, a Riemann hipotézishez kapcsolódó problémához, hogy numerikus alsó és felső korlátokat keressünk a de Bruijn–Newman-féle Λ konstansra. Ezen értekezés írásakor csak a Λ alsó korlátaira vonatkozó becslések ismeretesek; erről számolunk be az alábbiakban. E célból mindenekelőtt felsorolunk néhány hasznos elméleti eredményt az $\mathcal{L} - \mathcal{P}$ függvényekre vonatkozóan.

Először is a $z = -x^2$ helyettesítés (4.4)-ben az

$$(4.8) \quad F_\lambda(z) := \sum_{m=0}^{\infty} \frac{\hat{b}_m(\lambda) z^m}{(2m)!}$$

függvényt szolgáltatja, ahol $\hat{b}_m(\lambda)$ a (4.5) által adott. Mivel a (4.4)-beli $H_\lambda(x)$ és a (4.8)-beli $F_\lambda(z)$ kielégíti a

$$(4.9) \quad H_\lambda(x) = F_\lambda(-x^2) \quad (\lambda \in \mathbb{R})$$

relációt, azért $F_\lambda(z)$ nyilván egy $1/2$ rendű valós egész függvény, amelynek (lásd Boas [1, 24. oldal]) szükségképpen végtelen sok (esetleg néhány nem valós) gyöke van, a valós λ bármely választása mellett. Ezenkívül, mivel $\hat{b}_m(\lambda) > 0$ maga után vonja, hogy F_λ pozitív a $z \geq 0$ sugáron, a (4.7) and (4.9) alapján

$$(4.10) \quad \begin{cases} F_\lambda(z) & \text{csak valós negatív gyökökkel rendelkezik, ha } \lambda \geq \Lambda, \text{ és} \\ F_\lambda(z) & \text{rendelkezik nem valós gyökökkel, ha } \lambda < \Lambda. \end{cases}$$

Másrészt a (2.10)-beli Laguerre–Pólya függvényosztály definíciója szerint (4.10) tömören úgy fejezhető ki, hogy

$$(4.11) \quad \begin{cases} F_\lambda \in \mathcal{L} - \mathcal{P}, & \text{ha } \lambda \geq \Lambda, \text{ és} \\ F_\lambda \notin \mathcal{L} - \mathcal{P}, & \text{ha } \lambda < \Lambda. \end{cases}$$

Ezt a következőképpen használhatjuk fel. Tekintsünk egy tetszőleges $G(z) \in \mathcal{L} - \mathcal{P}$ elemet, amely a

$$(4.12) \quad G(z) = \sum_{m=0}^{\infty} \frac{\gamma_m}{m!} z^m$$

alakban írható, és legyen

$$(4.13) \quad G_n(t) := \sum_{k=0}^n \binom{n}{k} \gamma_k t^k \quad (n = 0, 1, \dots)$$

az ehhez tartozó n -edik Jensen polinom. Pólya és Schur [27] egy ismert eredménye szerint

$$(4.14) \quad G(z) \in \mathcal{L} - \mathcal{P} \iff G_n(t) \in \mathcal{L} - \mathcal{P} \quad (n = 0, 1, \dots),$$

másszóval

$$(4.15) \quad G(z) \in \mathcal{L} - \mathcal{P} \iff G_n(t) (\neq 0) \text{ csak valós gyökökkel rendelkezik, minden } n = 1, 2, \dots \text{ esetén.}$$

A (4.15) alapján azonnal kapjuk az alábbi eredményt.

6. ÁLLÍTÁS ([6]). Tegyük fel, hogy létezik egy pozitív egész m és egy valós λ úgy, hogy a (4.8)-beli $F_\lambda(z)$ m -edik Jensen polinomja, vagyis a

$$(4.16) \quad G_m(t; \lambda) := \sum_{k=0}^m \binom{m}{k} \frac{\hat{b}_k(\lambda) \cdot k!}{(2k)!} t^k$$

függvény rendelkezik egy nem valós gyökkel. Ekkor

$$(4.17) \quad \lambda < \Lambda.$$

Ez az állítás az alapja az alábbi, a de Bruijn–Newman-féle Λ konstans alsó korlátainak megkeresésére szolgáló numerikus algoritmusnak:

- (i) rögzítsünk egy valós $\lambda < 0$ értéket;
- (ii) (4.5) alapján határozzunk meg nagy pontosságú $\{\beta_k(\lambda)\}_{k=0}^N$ becsléseket a $\{\hat{b}_k(\lambda)\}_{k=0}^N$ momentumokra vonatkozóan;
- (iii) az (ii) és (4.16) alapján alkossuk meg a

$$(4.18) \quad g_m(t; \lambda) := \sum_{k=0}^m \binom{m}{k} \frac{\hat{\beta}_k(\lambda) k!}{(2k)!} t^k \quad (m = 1, 2, \dots, N)$$

approximatív Jensen polinomokat;

- (iv) egy nagy pontosságú gyökkereső algoritmussal keressük meg a $g_m(t; \lambda)$ ($m = 1, 2, \dots, N$) gyökeit;
- (v) amennyiben létezik egy pozitív egész m ($1 \leq m \leq N$) úgy, hogy a (iii)-beli $g_m(t; \lambda)$ m -edik Jensen polinomnak van egy nem valós gyöke, akkor elméleti úton bizonyítsuk be, hogy a megfelelő $G_m(t; \lambda)$ pontos Jensen polinom is *szükségszerűen* rendelkezik egy nem valós gyökkel. Ekkor $\lambda < \Lambda$.

Analízissel foglalkozó szakemberek számára a fent használt „szükségszerűen” kifejezés nyilván meglepően hangzik. Ezt mélyebben a következő részben elemezzük.

3.5. A Λ alsó korlátai a Jensen polinomok alapján

Csordas, Norfolk és Varga [6] arra használták a 6. Állítást, hogy egy bizonyos alsó korlátot keressenek a de Bruijn–Newman-féle Λ konstansra. Nevezetesen, a $\lambda = -50$ választással numerikusan kiszámíthatók a $\{\hat{b}_m(-50)\}_{m=0}^{16}$ momentumok, mégpedig az ún. *Romberg-féle integrálási módszerrel* (lásd Stoer és Bulirsch [31, 136. oldal]), legalább 60 értékes jegy relatív pontossággal, és a megfelelő (4.18)-beli 16-ik $g_{16}(t; -50)$ Jensen polinom egy nem valós gyöke

$$(5.1) \quad z_1 := -220,91911\,17368\,44951 \dots + i7,09256\,52553\,63889\,67 \dots,$$

a

$$(5.2) \quad |z_1| = 221,03293\,51307\,13450 \dots$$

abszolút értékkel. Annak bizonyítására, hogy a (4.16)-beli $G_{16}(t; -50)$ (pontos) Jensen polinom úgyszintén *garantáltan* rendelkezik egy nem valós gyökkel, [6]-ban felhasználtuk Ostrowski [23, Appendix B] alábbi eredményét.

7. ÁLLÍTÁS (Ostrowski [23]). Legyen $f(z) = \sum_{j=0}^n a_j z^j$ (ahol $a_0 a_n \neq 0$) és $h(z) = \sum_{j=0}^n b_j z^j$ két komplex polinom, és legyenek az $f(z)$ zérushelyei $\{z_j\}_{j=1}^n$ (ezek valamennyien zérustól különbözőek, de többszörös gyökök megengedettek). Tegyük fel, hogy létezik egy pozitív valós τ szám úgy, hogy $4n\tau^{1/n} \leq 1$ és

$$(5.3) \quad |b_j - a_j| \leq \tau a_j \quad (j = 0, 1, \dots, n).$$

Ekkor a $h(z)$ függvénynek a $\{w_j\}_{j=1}^n$ n zérushelye elrendezhető úgy, hogy

$$(5.4) \quad |w_j - z_j| < 8n\tau^{1/n} |z_j| \quad (j = 1, 2, \dots, n).$$

Ennek felhasználásával a következő eredményt kaptuk [6]-ban.

8. TÉTEL ([6]). Ha a de Bruijn–Newman-féle Λ konstans, akkor

$$(5.5) \quad -50 < \Lambda.$$

Bizonyítás. Legyen $f(z) := g_{16}(z; -50) = \sum_{k=0}^{16} \binom{16}{k} \frac{\hat{\beta}_k(-50) k! z^k}{(2k)!}$ és $h(z) := G_{16}(z; -50) = \sum_{k=0}^{16} \binom{16}{k} \frac{\hat{b}_k(-50) k! z^k}{(2k)!}$. Ekkor, mivel az approximált $\hat{\beta}_k(-50)$ és a pontos $b_k(-50)$ momentumok eltérésére teljesül, hogy

$$|\hat{b}_k(-50) - \hat{\beta}_k(-50)| \leq 10^{-60} \hat{\beta}_k(-50) \quad (k = 0, 1, \dots, 16),$$

láthatjuk, hogy (5.3) fennáll a $\hat{\tau} := 10^{-60}$ értékkel, s ezért $8n\hat{\tau}^{1/n} \leq 2.27620 \cdot 10^{-2}$, hacsak $n = 16$. Így az (5.4)-beli 7. Állítás az (5.1)-beli z_1 -re azt adja, hogy a $G_{16}(z; -50)$ függvénynek van egy w_1 gyöke úgy, hogy

$$(5.6) \quad |w_1 - z_1| \leq (2, 27620 \cdot 10^{-2}) |z_1| < 5, 03116.$$

Mivel $|\operatorname{Im} w_1 - \operatorname{Im} z_1| \leq |w_1 - z_1|$, a fenti egyenlőtlenség (5.1)-gyel együtt adja, hogy

$$|\operatorname{Im} w_1 - 7, 09256 \dots| < 5, 03116,$$

s innen

$$(5.7) \quad \operatorname{Im} w_1 > 2, 06164.$$

Ezért $G_{16}(z; -50)$ rendelkezik egy nem valós zérushellyel, és a 6. Állításból kapjuk a kívánt eredményt. \square

A Λ (5.5) alatti alsó korlátja volt az *első* ilyen korlát. Mint azt [6]-ban megemlítettük, ezen eredmény fontossága nem magában a konkrét alsó korlátban, hanem inkább a módszer *konstruktív voltában* rejlik. E dolgozatban ugyancsak említést nyert, hogy bizonyos optimizmussal tekinthetünk ezen alsó korlát javításának lehetőségére.

Kiderült, hogy az Ostrowski-féle 7. Állítás, amely *globális* relatív pontosságot szolgáltat az *összes* két közeli polinom gyökeinek eltérésére, túlságosan *konzervatív* a mi céljainkra abban a formában, ahogy a 8. Tétel bizonyításánál összehasonlítottuk két közeli polinom két speciális *egyszeres* gyökét. Ezzel kapcsolatban az alábbi elemi, de hasznos eredményt ismertetjük.

9. LEMMA. Legyen $p(z)$ egy komplex, n -edfokú polinom. Ha $p'(z_1) \neq 0$, akkor a

$$(5.8) \quad |z - z_1| \leq n|p(z_1)| / |p'(z_1)|$$

kör a $p(z)$ legalább egy zérushelyét tartalmazza.

Bizonyítás. Mivel a 9. Lemma eredménye triviális, ha $p(z_1) = 0$, tegyük fel, hogy $p(z_1) \neq 0$, és legyen $p(z) = \gamma \prod_{k=1}^n (z - \zeta_k)$, ahol ζ_k a $p(z)$ gyökei. A $p(z)$ logaritmikus deriváltját véve, és kiszámítva azt a z_1 pontban kapjuk, hogy

$$\frac{p'(z_1)}{p(z_1)} = \sum_{k=1}^n \frac{1}{z_1 - \zeta_k}.$$

Itt abszolút értékekre térve adódik, hogy

$$\frac{|p'(z_1)|}{|p(z_1)|} \leq \sum_{k=1}^n \frac{1}{|z_1 - \zeta_k|} \leq \frac{n}{\min_{1 \leq k \leq n} |z_1 - \zeta_k|},$$

ami nem más mint (5.8) más formában. \square

Megjegyezzük, hogy a 9. Lemma eredménye egy általánosabb eredmény speciális esete (lásd Henrici [14, 454. oldal]).

Anélkül, hogy a részletekbe mennénk, megemlíti, hogy ha a 7. Állítás helyett a 9. Lemmát használjuk, akkor a $\delta_m(\lambda)$ momentumok kiszámításánál sokkal *kevesebb* számjegyre van szükség annak garantálásához, hogy $G_{16}(z; -50)$ -nek nincs nem valós gyöke. E módszerrel Varga, Norfolk és Ruttan [33] szukcesszíve megjavították a Λ értékeit. (Lásd a 3.2. Táblázatot, ahol a második oszlop azon első Jensen polinom fokát tünteti fel, amelynél egy nem valós gyök *garantálható*.)

Ahogy λ növekedőleg zérus felé tart, a táblázatból láthatjuk, hogy a nem valós gyököt szolgáltató Jensen polinomok *foka* gyorsan növekszik. Hasonlóképpen gyorsan növekszik a *szükséges számjegyek száma* (harmadik oszlop). Ennek alátámasztására te Riele [29], a Jensen polinomos megközelítést alkalmazva, de azt Sturm sorozatok segítségével módosítva, nemrég bejelentette a

$$(5.9) \quad -5 < \Lambda$$

λ	fokszám	szükséges számjegyek	gyök
-100	10	12	$-45,840 \dots + i8,703 \dots$
-50	16	12	$-220,919 \dots + i7,092 \dots$
-20	41	18	$-111,065 \dots + i1,322 \dots$
-15	56	20	$-79,834 \dots + i0,282 \dots$
-10	97	21	$-45,530 \dots + i0,156 \dots$
-8	142	21	$-30,993 \dots + i0,124 \dots$

3.2. Táblázat

alsó korlátot, számításait egy 406-odfokú Jensen polinomra és 250 számjegy pontos-ságra alapozva! A következő részben egy másik eljárást ismertetünk, amellyel egy jelentősen javított alsó korlátot nyerünk a de Bruijn-Newman-féle Λ konstansra.

3.6. A $F_\lambda(z)$ gyökeiről

A (4.8) definícióból világos, hogy ha valamely valós $\lambda < 0$ esetén a (4.8)-beli $F_\lambda(z)$ rendelkezik nem valós gyökkel, akkor (4.11) alapján $F_\lambda(z) \notin \mathcal{L} - \mathcal{P}$; s ezért $\lambda < \Lambda$. Mivel $F_\lambda(z)$ egy $1/2$ rendű egész függvény, szükségszerűen végtelen sok gyöke van (lásd Boas [1, 24. oldal]). Mármint az ötlet az, hogy *közvetlenül nyomom követjük* az $F_\lambda(z)$ valamely gyökét. A továbbiak kedvéért megjegyezzük, hogy a (4.8)-beli $F_\lambda(z)$, a (4.3) analógiájára, kifejezhető az alábbi integrál formában:

$$(6.1) \quad F_\lambda(z) = \int_0^\infty e^{\lambda t^2} \Phi(t) \cosh(t\sqrt{z}) dt \quad (\lambda \in \mathbb{R}).$$

Most valamely valós λ_0 mellett legyen $z(\lambda_0)$ egy *egyszeres* gyöke az $F_{\lambda_0}(z)$ függvénynek, úgy, hogy $z(\lambda)$ egyszeres gyöke maradjon az $F_\lambda(z)$ függvénynek a λ valamely kis, λ_0 -t tartalmazó intervallumának belsejében. Ezen intervallumban $F_\lambda(z(\lambda)) \equiv 0$, úgyhogy a $\hat{b}_m(\lambda)$ (4.5)-beli definíciójával élve,

$$(6.2) \quad \begin{aligned} F_\lambda(z(\lambda)) \equiv 0 &= \int_0^\infty e^{\lambda t^2} \Phi(t) \cosh(t\sqrt{z(\lambda)}) dt \\ &= \sum_{m=0}^\infty \frac{\hat{b}_m(\lambda)}{(2m)!} (z(\lambda))^m. \end{aligned}$$

Differenciálva (6.2)-t a λ változóban, kapjuk, hogy

$$0 \equiv \sum_{m=0}^\infty \frac{\hat{b}_{m+1}(\lambda)(z(\lambda))^m}{(2m)!} + \frac{dz(\lambda)}{d\lambda} \sum_{m=0}^\infty \frac{(m+1)\hat{b}_{m+1}(\lambda)(z(\lambda))^m}{(2m+2)!}.$$

Mivel a fenti, $dz(\lambda)/d\lambda$ szorzójaként fellépő összeg zérustól különböző (hiszen a feltevés szerint $z(\lambda)$ egy egyszeres gyök), azért innen

$$(6.3) \quad \frac{dz(\lambda)}{d\lambda} = - \frac{\sum_{m=0}^\infty \hat{b}_{m+1}(\lambda)(z(\lambda))^m / (2m)!}{\sum_{m=0}^\infty (m+1)\hat{b}_{m+1}(\lambda)(z(\lambda))^m / (2m+2)!}.$$

Úgysszintén fontos megjegyezni, hogy mivel $\hat{b}_m(\lambda) := \int_0^\infty t^{2m} e^{\lambda t^2} \Phi(t) dt$, ezért a $e^{\lambda t^2}$ függvényt helyettesítve Maclaurin sorával és tagonként integrálva, adódik

$$(6.4) \quad \hat{b}_m(\lambda) = \sum_{j=0}^{\infty} \frac{\hat{b}_{m+j}(0) \lambda^j}{j!} \quad (j = 0, 1, \dots; \lambda \in \mathbb{R}).$$

Így kis negatív λ mellett (6.4)-ből a $\{\hat{b}_m(0)\}_{m=0}^N$ nagy pontosságú momentumoknak *csak egy kiterjesztett táblázatát* szükséges kiszámítani, s ebből a $\{\hat{b}_m(\lambda)\}_{m=0}^{N'}$ momentumok (6.4) segítségével közvetlenül becsülhetők. Valóban, a Kent State University matematikai tanszékén az Encore paralell számítógépen kiszámítottuk a $\{\hat{b}_m(0)\}_{m=0}^{1600}$ momentumokat, valamennyit 220 értékes jegyre, egyszerűen a trapéz szabály megfelelően finom felosztás melletti alkalmazásával. Pontosabban, ha $T_h(m)$ jelöli a $\hat{b}_m(0)$ trapéz szabály szerinti approximációját (egy h nagyságú beosztással), akkor (2.5ii) alapján megmutatható (lásd [33]), hogy

$$|T_h(m) - \hat{b}_m(0)| = O\left(\exp\left(\frac{-(\pi^2 - \varepsilon)}{4h}\right)\right) \quad (h \rightarrow 0).$$

Ez a geometriai nagyságrendű konvergencia teszi a trapéz szabály alkalmazását gyorsá és pontosá.

A továbbiakban jelölje $\{\rho_j := \frac{1}{2} + it_j\}_{j=1}^{15\,000}$ a ζ -függvény ismert egyszeres gyökeit a $\operatorname{Re} z = 1/2$ kritikus egyenesen. Ezeket te Riele [28] egy táblázatban foglalta össze. A $z = -x^2$ változó-transzformációval adódik, hogy $\{z_j(0) := -4t_j^2\}_{j=1}^{15\,000}$ gyökei az $F_0(z)$ függvénynek. Kiderült, hogy az $F_0(z)$ egymásutáni gyökeiből alkotott párok *némelyike* egészen közel áll egymáshoz. Egy ilyen kiragadott pár

$$(6.5) \quad z_{34}(0) = -49\,310,231\dots, \text{ és } z_{35}(0) = -50\,063,757\dots,$$

és a (6.3) formulával kiszámítottuk, hogy

$$(6.6) \quad \left. \frac{dz_{34}(\lambda)}{d\lambda} \right|_{\lambda=0} = +877,836\dots, \text{ és } \left. \frac{dz_{35}(\lambda)}{d\lambda} \right|_{\lambda=0} = -26,627\dots$$

E két mennyiség ellenkező előjele azt jelenti, hogy a két gyök *egymáshoz tart*, midőn λ zérusra csökken, vagyis e gyökök *vonzzák* egymást.

Az $F_\lambda(z)$ gyökpárjainak e nyomkövetése néhány érdekes geometriai eredményt produkál. A 3.1. ábrán a

$$\{z_{34}(-[0.04]j) \text{ és } z_{35}(-[0.04]j)\}_{j=0}^{20}$$

21 gyökpárt jelenítettük meg. Láthatjuk, hogy a $z_{34}(\lambda)$ és $z_{35}(\lambda)$ valós gyökként indulnak, és egymás felé tartanak, majd *találkoznak*, az $F_\lambda(z)$ egy *kétszeres gyökét* alkotva, midőn $\lambda \doteq 0.38$. Ezekután e gyökök egy nem valós konjugált komplex párba ágaznak szét, s egy parabola-szerű trajektóriát írnak le a komplex síkon. A

3.1. ábráról azt is látjuk, hogy a gyökpár, $z_{34}(\lambda)$ és $z_{35}(\lambda)$, két *nem valós* gyököt szolgáltat az $F_\lambda(z)$ függvényre nézve, ha csak $\lambda \leq -0.40$. A λ ezen értékeire $F_\lambda(z)$ nem lehet az $\mathcal{L} - \mathcal{P}$ függvényosztály tagja, és (4.11)-ből következik, hogy -0.40 egy alsó korlát a de Bruijn–Newman-féle Λ konstansra. Valóban, egy, a 9. Lemmára alapozott, meglehetősen hosszú analízissel egy jobb alsó korlát garantálható. Ezzel kapcsolatban Varga, Norfolk és Ruttan [33] eredménye a következőképpen fogalmazható meg:

10. TÉTEL ([33]). Ha Λ jelöli a de Bruijn–Newman konstans, akkor

$$(6.7) \quad -0.385 < \Lambda.$$

A fenti módszer egy jelentős numerikus hátránya a de Bruijn–Newman féle Λ konstans alsó korlátainak meghatározásánál a (6.1)-beli $F_\lambda(z)$ valós egész függvény *nem valós gyökei* nagypontosságú nyomkövetésének igényéből fakad. Számítástechnikai szempontból nyilván egyszerűbb lenne az összes nagypontosságú számítást kizárólag *valós* aritmetikában végezni. Nemrég Csordas, Ruttan és Varga [34] egy új módszert találtak a Λ alsó korlátainak meghatározására, amelyben ez a szempont érvényesül. Az ötlet meglehetősen egyszerű. A (4.3)-beli $H_\lambda(x)$ valós egész függvényre, ahol λ valós, (4.9) alapján a (4.11) alábbi analogonja mondható ki:

$$(6.8) \quad \begin{cases} H_\lambda \in \mathcal{L} - \mathcal{P}, & \text{ha } \lambda \geq \Lambda, \text{ és} \\ H_\lambda \notin \mathcal{L} - \mathcal{P}, & \text{ha } \lambda < \Lambda. \end{cases}$$

Így bármely $\lambda \geq \Lambda$ esetén (6.8)-ból és a 2. Tétel (2.11)-beli $p = 0$ esetéből szükségszerűen következik, hogy

$$(6.9) \quad L_1(H_\lambda(x)) := (H'_\lambda(x))^2 - H_\lambda(x) \cdot H''_\lambda(x) \geq 0 \quad (x \in \mathbb{R}).$$

Innen, ha egy valós $\hat{\lambda}$ és egy valós X található úgy, hogy $L_1(H_{\hat{\lambda}}(X)) < 0$, akkor $\hat{\lambda}$ automatikusan egy alsó korlát a Λ -ra, vagyis

$$\hat{\lambda} < \Lambda.$$

A (6.9)-cel kapcsolatban megjegyezzük, hogy az $L_1(H_\lambda(x))$ kiszámítása csak a (4.3) által adott

$$\int_0^\infty e^{\lambda t^2} \Phi(t) \cos(xt) dt, \quad -\int_0^\infty t e^{\lambda t^2} \Phi(t) \sin(xt) dt, \quad \text{és} \quad -\int_0^\infty t^2 e^{\lambda t^2} \Phi(t) \cos(xt) dt.$$

valós integrálok nagypontosságú kiszámítását igényli.

Speciálisan, ezen eljárás alkalmazásával a 10. Tétel alábbi javítása kapható nagypontosságú valós aritmetikával:

11. TÉTEL ([34]). Ha λ a de Bruijn–Newman konstans, akkor

$$(6.10) \quad -0.0991 < \Lambda.$$

3.7. Szükséges és elégséges feltételek a Riemann hipotézisre

E fejezetet a Riemann hipotézisre vonatkozó néhány ismert szükséges és elégséges feltétel tárgyalásával fejezzük be. Először is egy ismert, lényegében Jentsentől [16] származó eredményt fogalmazunk meg. (Ennek egy teljes bizonyítása megtalálható Csordas és Varga [8] dolgozatában.)

12. TÉTEL (Jensen [16]). Legyen

$$(7.1) \quad f(z) = e^{-\alpha z^2} f_1(z) \quad (\alpha \geq 0, f(z) \not\equiv 0),$$

ahol f_1 egy 0 vagy 1 génuszú valós egész függvény. Ekkor $f \in \mathcal{L} - \mathcal{P}$ akkor és csak akkor, ha

$$(7.2) \quad |f'(z)|^2 \geq \operatorname{Re} \{f(z)\overline{f''(z)}\} \quad (z \in \mathbb{C}).$$

A (2.6)-ból és a (4.3)-beli $\lambda = 0$ eset tárgyalásából következik, hogy

$$(7.3) \quad H_0(x) = \xi(x/2)/8 = \frac{1}{2} \int_{-\infty}^{+\infty} \Phi(t) e^{ixt} dt,$$

ahol $\Phi(t)$ rendelkezik a (2.5) tulajdonságokkal. Mint láttuk, a Riemann hipotézis akkor és csak akkor igaz, ha a $H_0(x)$ összes gyöke valós, ami úgyszintén ekvivalens azzal az állítással, hogy $H_0 \in \mathcal{L} - \mathcal{P}$. Ily módon közvetlenül alkalmazva a (7.2) feltételt a (7.3)-beli $H_0(z)$ függvényre (s ezzel kettős integrálokat teremtve), Jensen alábbi ismert (publikálatlan) eredményét kapjuk, amely Pólya [25]-ben jelent meg.

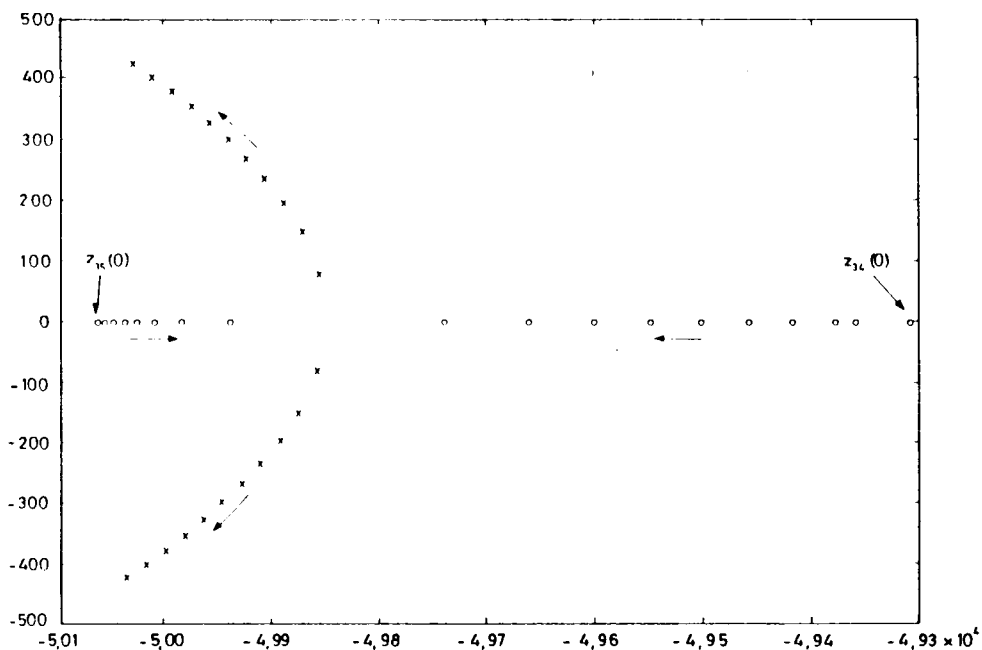
13. TÉTEL (Pólya [25]). A Riemann hipotézis fennállásának egy szükséges és elégséges feltétele az, hogy

$$(7.4) \quad \Delta(x, y) \geq 0, \text{ ha csak } x, y \in \mathbb{R},$$

ahol

$$(7.5) \quad \Delta(x, y) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(t) \Phi(s) e^{i(t+s)x} e^{(t-s)y} (t-s)^2 dt ds.$$

Szimmetria okokból (lásd (2.5iii)), elegendő kiszámítani a (7.5)-beli $\Delta(x, y)$ értékeket csupán $x, y \geq 0$ esetén, de a (7.4.) alatti $\Delta(x, y) \geq 0$ feltételek ellenőrzése numerikus integrálással még ezzel a redukcióval is félelmetes feladatnak látszik. Mindazonáltal Csordas és Varga [8] egy újabb eredménye, amely a Hermite–Biehler tétel kiterjesztéseit használja, tovább redukálja azt a tartományt, ahol a $\Delta(x, y)$ pozitivitását kell bizonyítani.



9.1. ábra $\{z_{34}(-[0, 04]j)\}_{j=0}^{20}$ és $\{z_{35}(-[0, 04]j)\}_{j=0}^{20}$

14. TÉTEL ([8]). A Riemann hipotézis fennállásának egy szükséges és elégséges feltétele, hogy

$$(7.6) \quad \Delta(x, y) \geq 0 \text{ bármely } 0 < x < \infty \text{ és } 0 \leq y < 1 \text{ esetén.}$$

Más hasonló, kettős integrálokat tartalmazó szükséges és elégséges feltételek a Riemann hipotézis fennállására megtalálhatók Pólya [25], Csordas és Varga [9], és legújabban Csordas, Smith és Varga [35] munkáiban. Ez utóbbiban például bebizonyítjuk, hogy a (7.6) feltétel tovább redukálható a következőre:

$$(7.7) \quad \Delta(x, y) > 0 \text{ ha csak } x > 10^9 \text{ és } 0 < y < 1.$$

IRODALOM

- [1] BOAS, R. P., *Entire Functions* (Academic Press, Inc., New York, 1954).
- [2] DE BRUIJN, N. G., „The roots of trigonometric integrals”, *Duke J. Math.* 17 (1950), 197–226.
- [3] CONREY, J. B., „At least two fifths of the zeros of the Riemann zeta function are on the critical line”, *Bull. Amer. Math. Soc.* 20 (1989), 79–81.
- [4] CSORDAS, G., NORFOLK, T. S. és VARGA, R. S., „The Riemann Hypothesis and the Turán inequalities”, *Trans. Amer. Math. Soc.* 296 (1986), 521–541.
- [5] CSORDAS, G., NORFOLK, T. S. és VARGA, R. S., *publikálatlan kézirat.*
- [6] CSORDAS, G., NORFOLK, T. S. és VARGA, R. S., „A lower bound for the de Bruijn-Newman constant Λ ”, *Numer. Math.* 52 (1988), 483–497.
- [7] CSORDAS, G. és VARGA, R. S., „Moment inequalities and the Riemann Hypothesis”, *Constr. Approx.* 4 (1988), 175–198.
- [8] CSORDAS, G. és VARGA, R. S., „Fourier transforms and the Hermite-Biehler Theorem”, *Proc. Amer. Math. Soc.* 107 (1989), 645–652.
- [9] CSORDAS, G. és VARGA, R. S., „Necessary and sufficient conditions and the Riemann Hypothesis”, *Adv. in Appl. Math.* 11 (1990), 328–357.
- [10] GROSSWALD, E., „Generalization of a formula of Hayman, and its applications to the study of Riemann's zeta function”, *Illinois J. Math.* 10 (1966), 9–23.
- [11] GROSSWALD, E., „Correction and completion of the paper 'Generalization of a formula of Hayman'”, *Illinois J. Math.* 13 (1969), 276–280.
- [12] HARDY, G. H., „Sur les zéros de la fonction $\zeta(s)$ de Riemann”, *C.R. Acad. Sci. Paris* 158 (1914), 1012–1014.
- [13] HAYMAN, W. K., „A generalization of Stirling's formula”, *J. Reine Angew. Math.* 196 (1956), 67–95.
- [14] HENRICI, P., *Applied and Computational Complex Analysis*, vol. 1 (John Wiley & Sons, New York, 1974).
- [15] HENRICI, P., *Applied and Computational Complex Analysis*, vol. 2 (John Wiley & Sons, New York, 1977).
- [16] JENSEN, J. L. W. V., „Recherches sur la théorie des équations”, *Acta Math.* 36 (1913), 181–195.
- [17] LAGUERRE, E., *Oeuvres*, vol. 1 (Gauthier-Villars, Paris, 1898).
- [18] LEVINSON, N., „More than one third of the zeros of Riemann's zeta-function are on $\sigma = 1/2$ ”, *Adv. in Math.* 13 (1974), 383–436.
- [19] VAN DE LUNE, J., TE RIELE, H. J. J. és WINTER, D. T., „On the zeros of the Riemann zeta function in the critical strip IV”, *Math. Comp.* 46 (1986), 667–681.
- [20] MATIYASEVICH, YU. V., „Yet another machine experiment in support of Riemann's conjecture”, *Kibernetika* (1982), (Oroszul, angol nyelvé kivonattal); *Angol fordítás: Cybernetics* 18 (1983), 705–707.
- [21] NEWMAN, C. M., „Fourier transforms with only real zeros”, *Proc. Amer. Math. Soc.* 61 (1976), 245–251.
- [22] ODLYZKO, A. M., *The 10^{20} -th zero of the Riemann zeta function and its neighbors* (1989), preprint.
- [23] OSTROWSKI, A. M., *Solution of Equations and Systems of Equations* (Academic Press, Inc., New York, 1960).
- [24] PÓLYA, G., „On the zeros of certain trigonometric integrals”, *J. London Math. Soc.* 1 (1926), 98–99.
- [25] PÓLYA, G., „Über die algebraisch-funktionentheoretischen Untersuchungen von J.L.W.V. Jensen”, *Kgl. Danske Vid. Sel. Math.-Fys. Medd.* 7 (1927), 3–33.
- [26] PÓLYA, G., „Über trigonometrische Integrale mit nur reellen Nullstellen”, *J. Reine Angew. Math.* 158 (1927), 6–18.
- [27] PÓLYA, G. és SCHUR, J., „Über zwei Arten von Faktorenfolgen in der Theorie der algebraischen Gleichungen”, *J. Reine Angew. Math.* 144 (1914), 89–113.

- [28] TE RIELE, H. J. J., „Table of the First 15,000 Zeros of the Riemann Zeta Function to 28 Significant Digits and Related Quantities”, *Report Number 67/79 of the Mathematisch Centrum, Amsterdam* (1979).
- [29] TE RIELE, H., „A new lower bound for the de Bruijn-Newman constant”, *Numer. Math.* **58** (1991), 661–667.
- [30] RIEMANN, B., „Über die Anzahl der Primzahlen unter einer gegebenen Grösse”, *Monatsh. der Berliner Akad.*, (1858/60), 671–680 (1982 No VII), 145–153, *Gesammelte Mathematische Werke*, 2nd edition (Teubner, Leipzig).
- [31] STOER, J. E'S BULIRSCH, R., *Introduction to Numerical Analysis* (Springer-Verlag, Heidelberg, 1980).
- [32] TITCHMARSH, E. C., *The Theory of the Riemann Zeta-function*, 2nd Edition (revised by D.R. Heath-Brown) (Oxford University Press, Oxford, 1986).
- [33] VARGA, R. S., NORFOLK, T. S. és RUTTAN, A., „A lower bound for the de Bruijn-Newman constant. II”, *Conference Proceedings, US-USSR Conference on Approximation Theory* (Tampa, 1990), (megjelenés alatt).
- [34] CSORDAS, G., RUTTAN, R. és VARGA, R. S., *Numerical Algorithms* **1** (1991), 305–330.
- [35] CSORDAS, G., SMITH, W. és VARGA, R. S., „Level sets of real entire functions and the Laguerre inequalities”, *Trans. Amer. Math. Soc.*, (megjelenés alatt).

RICHARD S. VARGA
INSTITUTE FOR COMPUTATIONAL MATHEMATICS
KENT STATE UNIVERSITY
KENT, OH 44242

Fordította: Szabados József

KÖNYVISMERTETÉS

V. LAKSHMIKANTHAM, S. LEELA AND A. A. MARTYNYUK, : *Stability Analysis of Nonlinear Systems* (Pure and Applied Mathematics, vol. 125), Marcel Dekker, New York – Basel, 1989.

Ez a monográfia egy alapos, rendszerezett áttekintést ad néhány olyan technikáról, amellyel az utóbbi évtizedekben Ljapunov direkt módszerét tökéletesítették.

Ljapunov éppen most száz éve közölte azóta híressé vált doktori értekezését, amellyel meg-alapozta az egzakt stabilitáselméletét, mint matematikai diszciplinát. Ez a munka még ma is a stabilitáselmélet irányt szabó, alapvető monográfiája; a benne ismertetett direkt módszer ma is a legelterjedtebb, leghatékonyabb. Ez elsősorban annak köszönhető, hogy alkalmazásához nincs szükség a megoldások ismeretére. Ljapunov zseniális felismerése az volt, hogy a mechanikai rendszerek stabilitásvizsgálatában oly' sikeres teljes mechanikai energia általánosítható minden differenciálegyenletekkel leírt rendszerre. Ha létezik bizonyos tulajdonságú skalárértékű függvény a rendszerhez, akkor a kitüntetett megoldás egy megfelelő stabilitási tulajdonsággal rendelkezik. Ez a módszer igen nagy lendületet kapott az ötvenes években, amikor Massera bebizonyította, hogy az alaptételek megfordíthatók: a stabilitási tulajdonság fennállásából következik a kívánt tulajdonságú függvény létezése. Bebizonyosodott, hogy a módszer sikeresen használható nemcsak közönséges, de funkcionáldifferenciálegyenletek és parciális differenciálegyenletek által leírt rendszerekre, továbbá az irányításelméletben is.

A jelen monográfia arról szól, hogy milyen kapcsolat van a száz éves direkt módszer és a differenciálegyenlőtlenségek jóval fiatalabb, negyven éves elmélete között, és hogy ez a kapcsolat hogyan használható a direkt módszer tökéletesítésére, kiterjesztésére. A tételeket igen általánossá teszi az, hogy két különböző mértékre alapozott stabilitási fogalmakról szólnak. Ez durván szólva annyit jelent, hogy a kezdeti értékek perturbációját más mértékkel mérjük, mint a megoldások eltérését. Ennek köszönhetően az eredmények magukban foglalnak különböző stabilitási fogalmakat (például, a parciális stabilitást.)

A könyv értékei között kiemelendő, hogy áttekintést ad az állandók variálása módszerének nem-lineáris kiterjesztéséről, annak alkalmazásairól, és a LakshmiKantham által kidolgozott monoton iterációs technikáról. A módszereket különböző alkalmazási területekről (mechanika, populációdinamika, közgazdaságtan, rugalmasságtan, reakciókinetika) választott példákkal illusztrálják.

Melegen ajánlom ezt a monográfiát mind a matematikusok, mind a matematika alkalmazói számára. Segítségével nemcsak áttekintést nyerhet az olvasó a nem-lineáris rendszerek stabilitáselméletének legújabb eredményeiről, de egyben inspirációt kaphat a kutatásokba való bekapcsolódásra is.

Hatvani László (Szeged)

A kiadásért felelős az ELTE TTK dékánja
Szedte a KLTE Informatikai és Számítóközpont Kiadvány Szerkesztő Csoportja
és nyomta az ELTE Sokszorosítóüzeme
Felelős vezető: Arató Tamás
Budapest, 1992. – ELTE 94014
Megjelent: 19,25 (A/5) ív terjedelemben
350 példányban
HU ISSN 0133-3399

ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését, olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartalmazzon. A közlésre szánt dolgozatokat három példányban kell beküldeni. Előnyben részesülnek a TEX-ben elkészített dolgozatok. Ezeket két kinyomtatott példány kíséretében diszketten kérjük beadni.

A kéziratok szerkezeti felépítésének a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámmal kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék mindig az utolsó szakasz kell, hogy legyen, és azt nem kell sorszámmal ellátni. Az irodalomjegyzék után, a kézirat befejezésekképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos postai címét. A dolgozatban előforduló képleteket szakaszonként újrakezdődően, a képlet előtt két zárójel közé írt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segéditételeket és lemmákat) ugyan-csak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német, francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót. Amennyiben lehetséges, kérjük a nyomtatás számára különösen nehézkes matematikai jelölések használatának az elkerülését.

A dolgozatok ábráit és az esetleges lábjegyzeteket a dolgozat végén, különálló lapokon kérjük beküldeni. Mind az ábrákat, mind a lábjegyzeteket a dolgozat szakaszokra bontásától független, folytatódó arab sorszámozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, széljegyzetként feltüntetett, ábraazonosító sorszámokkal kell megadni. A lábjegyzetekre a dolgozaton belül az azonosító sorszám felső indexkénti használatával lehet hivatkozni.

Az irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a című szerzők nevét a Mathematical Reviews átirási szabályai szerint latin betűsre kell átírni. A folyóiratban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gépi program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., Über die Theorie der einfachen Ungleichungen, *Journal für die reine und angewandte Mathematik* 124 (1902) 1-27.
- [2] Kéri, G., „DUALSIMP”, rutin a CDC 3300-as gépekre (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertető 2. 1973. május) 19-20.
- [3] Prékopa, A., „Sztochasztikus rendszerek optimalizálási problémáiról”, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U., „Recent research on the ruin problem of collective risk theory”, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam-London, 1973) 221-228.
- [5] Zoutendijk, G., *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számai zárójelben kell megadni, mint például [5] vagy [4, 76-78]. A szerzők a dolgozatokról 50 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

TARTALOMJEGYZÉK

<i>Domokos Gábor</i> , Síkbeli, konzervatív tenzormezők forgásszimmetriáinak vizsgálata elemi módszerekkel	1
<i>Szemők Árpád</i> , Optimális ragadozási stratégia a ragadozó-zsákmány viszonyban	27
<i>Lengyel Imre</i> , Megjegyzések a Gilmore–Gomory módszer gyakorlati alkalmazásához	45
<i>Gregorics Tibor</i> , Megjegyzések a heurisztikus gráfkereső algoritmusokhoz	61
<i>Deák Jenő, Tóth János és Vizvári Béla</i> , Anyagmegmaradás összetett kémiai mechanizmusokban	73
<i>Lámer Géza</i> , A szükséges és elégséges összeférhetőségi peremfeltételek meghatározása	99
<i>Tusnádý Gábor</i> , A rákkutatás matematikai alapjai	115
<i>A külföldi szakirodalomból</i>	
<i>Saari, D. G.</i> , Káosz és a választások elmélete	131
<i>Varga, R. S.</i> , Matematikai problémák és sejtések numerikus megközelítésben	143
Könyvismertetés	215

INDEX

<i>Domokos, G.</i> , Investigation of the rotational symmetries of planar, conservative tensor fields by elementary methods	1
<i>Szemők, Á.</i> , Optimal predator functional response	27
<i>Lengyel, I.</i> , Computational notes to the Gilmore–Gomory cutting stock algorithm	45
<i>Gregorics, T.</i> , Remarks to the heuristic graph-search algorithms	61
<i>Deák, J., Tóth, J. and Vizvári, B.</i> , Conservation of matter in complex chemical mechanisms	73
<i>Lámer, G.</i> , Determination of necessary and sufficient compatibility boundary conditions ...	99
<i>Tusnádi, G.</i> , Mathematical methods of cancer research	115
<i>From the foreign literature</i>	
<i>Saari, D. G.</i> , Chaos and the theory of elections	131
<i>Varga, R. S.</i> , Scientific Computation on Mathematical Problems and Conjectures	143
Book review	215

317471

Alkalmazott matematikai lapok

1992/3-4

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK
OSZTÁLYÁNAK KÖZLEMÉNYEI

16.

KÖTET

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA MATEMATIKAI ÉS FIZIKAI TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

BENCZÜR ANDRÁS

FŐSZERKESZTŐ-HELYETTESEK

DEMETROVICS JÁNOS, FARKAS MIKLÓS

FELELŐS SZERKESZTŐ

SZÁNTAI TAMÁS

A SZERKESZTŐBIZOTTSÁG TAGJAI

Arató Mátyás, Csirik János, Csiszár Imre, Galántai Aurél, Gécseg Ferenc, Gyires Béla, Gyórfy László, Harnos Zsolt, Hatvani László, Heppes Aladár, Kátai Imre, Katona Gyula, Kis Ottó, Klafszky Emil, Kovács Margit, Lovász László, Maros István, Prékopa András, Recski András, Stoyan Gisbert, Tandori Károly, Tusnády Gábor, Varga László

XVI. kötet 3-4. szám

Szerkesztőség és kiadóhivatal: 1117 Budapest, Bogdánfy út 10/B.

Az Alkalmazott Matematikai Lapok változó terjedelmű füzetekben jelenik meg, és olyan eredeti tudományos cikkeket publikál, amelyek a gyakorlatban, vagy más tudományokban közvetlenül felhasználható új matematikai eredményt tartalmaznak, illetve már ismert, de színvonalas matematikai apparátus új szerző és jelentős alkalmazását mutatják be. A folyóirat közöl cikk formájában megírt, új tudományos eredménynek számító programokat, és olyan, külföldi folyóiratban már publikált dolgozatokat, amelyek magyar nyelven történő megjelentetése elősegítheti az elért eredmények minél előbbi, széles körű hazai felhasználását. A szerkesztőbizottság bizonyos időnként lehetővé kívánja tenni, hogy a legjobb cikkek nemzetközi folyóiratok különszámaként angol nyelven is megjelenhessenek.

A folyóirat feladata a Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztályának munkájára vonatkozó közlemények, könyvismertetések stb. publikálása is.

A kéziratok a főszerkesztőhöz, vagy a szerkesztőbizottság bármely tagjához beküldhetők. A főszerkesztő címe:

Benczúr András, főszerkesztő
1117 Budapest, Bogdánfy út 10/B.

Közlésre el nem fogadott kéziratokat a szerkesztőség lehetőleg visszajuttat a szerzőhöz, de a beküldött kéziratok megőrzéséért vagy továbbításáért felelősséget nem vállal.

Az Alkalmazott Matematikai Lapok előfizetési ára kötetenként 850 forint. Megrendelések a szerkesztőség címén lehetségesek (átutalásokat az ELTE MNB 232-90142-6207 számlaszámára a 9015567 munkaszám és 546 utalványozási kód megjelöléssel kérjük).

A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztálya a következő idegen nyelvű folyóiratokat adja ki:

1. Acta Mathematica Hungaricae,
2. Acta Physica Hungaricae,
3. Studia Scientiarum Mathematicarum Hungarica.

MATEMATIKAI PROGRAMOZÁSI SZAKÉRTŐI RENDSZEREKRŐL¹

BÍRÓ M., MAYER J., RAPCSÁK T. ÉS VERMES M.

MTA SZTAKI

A dolgozat tárgya annak vizsgálata, milyen követelményeket kell a matematikai programozási modellezési és modellkezelési rendszereknek teljesíteni ahhoz, hogy matematikai programozási szakértői rendszereket nyerjünk, amelyek intelligens eszközként szolgálhatnak matematikai programozási feladatok megoldására önállóan vagy más rendszerbe építve.

Az MTA SZTAKI Operációkutatási Osztályán egy matematikai programozási modellkezelő rendszer lett kifejlesztve, amelynek modellezési nyelve a GAMS és amely a matematikai programozási taxonómia egy megvalósítását is tartalmazza.

1. Bevezetés

Intelligens matematikai programozási eszközrendszerek kifejlesztése intenzív kutatások tárgya. A jelenlegi stádiumban a fő cél az alapelvek kikristályosítása, illetve az ezeknek az elveknek leginkább megfelelő részrendszerek tervezése és megvalósítása, ld. Geoffrion (1987) és Maturana (1987).

A bevezetésben először ismertetjük a döntéstámogatás folyamatában azokat a legfontosabb feladatokat, amelyek megoldása vagy megoldásának a segítése részét kell hogy képezze egy modellezési rendszer szolgáltatásainak. A problémák sokrétűsége mutatja, mi az oka annak, hogy az eddig implementált rendszerek hatóköre általában a megoldandó feladatok csak egy egészen szűk körére korlátozódik.

A döntéshozatali problémák központi kérdése a döntési helyzetek felismerése és modellezése. Ez az adott szakterület és a rendelkezésre álló megoldási módszerek alapos ismeretét tételezi fel. Például egy bonyolult termelési-gazdálkodási szervezet vezetési tevékenységeinek egyetlen modellben történő leírása és összefüggő folyamatként való irányítása ma még nem lehetséges, pedig már egy kisebb egységen, egy üzemben belül is a gazdasági, a műszaki, a környezeti tényezők bonyolult kölcsönhatása gyakran vet fel komplex döntéshozatali problémákat. Teljességükben az ilyen feladatok rendszerint nem oldhatók meg, azonban részekre, komponensekre bontva mindig több olyan döntési helyzetre bukkanunk, amelyek esetleg különböző típusú modellekbe beépíthetők és a számítástechnika felhasználásával megoldhatók.

Az operációkutatási modellalkotást és a feladatmegoldást a következő lépések jellemzik (Geoffrion, 1987 és 1988) és (Rapcsák, 1988):

¹A kutatás az OMFB-EKPI G1-16-138 számú, az OTKA-1200 és az OTKA-2568 számú szerződések keretében folyt.

1. *A döntési probléma verbális megfogalmazása*

A problémafelismerés folyamata a rendszerelemzésen keresztül vezet a probléma verbális megfogalmazásáig és a döntéshozatalhoz szükséges információk megnevezéséig.

2. *Modellezés*

A problémafelvetés, a megoldó algoritmusok és az azzal kapcsolatos számítógépes lehetőségek ismeretében történik a modellépítés. Ez általában a verbálisan megfogalmazott probléma egy matematikai reprezentációja.

3. *A modell információs rendszerének a megadása*

A probléma és a modell ismeretében kerülhet sor a modell információs rendszerének a kialakítására, amelynek a leglényegesebb elemei a modellre vonatkozó tudás és a modellhez kapcsolódó numerikus adatok.

4. *A megoldó algoritmus kiválasztása*

A modell és a modell információs rendszerének az ismeretében kerülhet sor a megoldó algoritmus kiválasztására (a program megírására), ami az esetek nagy részében csak egy része a valódi munkának, mert még a széleskörűen használt programok esetén is gyakran szükség van ezeknek a probléma sajátosságainak megfelelő módosítására.

5. *A modellből adódó feladatok megoldása*

A döntéstervezés és döntéselemzés segítése a modellből adódó feladatok megoldásával kezdődik. Ehhez szükség van a probléma egy újabb, az úgynevezett numerikus reprezentációjára.

6. *Az eredmények verifikálása*

Mivel egy modellkonstrukció során szükségszerűen bizonyos egyszerűsítések, elhanyagolások történnek, másrészt a megoldó algoritmusok sok esetben nem kellően veszik figyelembe a feladatok sajátosságait, ezért minden esetben célszerű a modellből nyert eredményeket ellenőrizni. Ezek a vizsgálatok az esetek nagy részében vagy a modell és a feladat, vagy a megoldó algoritmus, vagy mindhárom módosításához vezetnek. Gyakorlati problémák esetén csak néhány ilyen jellegű változtatás, továbbfejlesztés után várható kielégítő eredmény.

7. *Az eredmények alkalmazása*

Az eredmények ellenőrzése után kerülhet sor az alkalmazásra, amelyben a gyakorlati területen dolgozó vezetőké és a nagy tárgy- és helyismerettel rendelkező szakembereké a meghatározó szerep. Ennek előfeltétele az eredmények megfelelő reprezentációja, illetve az ehhez kapcsolódó indoklás, amely a felhasználó számára megkönnyíti az eredményhez vezető út megértését is.

8. A rendszer karbantartása és továbbfejlesztése

Az előbbi pontokban ismertetett követelményeknek eleget tevő alkalmazási rendszerek felhasználása esetén nagyon lényeges a folyamatos karbantartásuk és a továbbfejlesztésük, mivel minden területen szinte napról napra változó körülmények között folyik a termelés. Ennek következtében egyrészt az alkalmazási rendszerek adatstruktúrái és egyéb részrendszerei megváltozhatnak, másrészt a gyártási technológiák és a szoftverrendszerek erkölcsi kopása nagyon nagy lehet. Továbbá lehetővé kell tenni, hogy a rendszer működése során nyert értékes információk és tapasztalatok a rendszerbe könnyen beépíthetők legyenek, amivel a hatékonyság növelhető.

Ezek a részfeladatok minden döntési probléma megoldására jellemzőek, ezért a matematikai programozási szakértői rendszereknek is mindezen szempontok figyelembevételével kell készülnie. Látható, hogy a matematikai programozási szakértői rendszerek készítésénél a súlyponti kérdés a különböző reprezentációk (verbális, matematikai, tudás, numerikus, eredmény) ábrázolása és ezek interakcióinak megvalósítása (Sprague és Carlson, 1982), (Binbasioglu és Jarke, 1985), (Dolk és Konsynski, 1985), (Dolk, 1986) és (Bíró, 1989).

A dolgozat bevezetőjében röviden áttekintettük azokat az általános szempontokat, melyeket minden modellezési rendszer felépítésénél, így az intelligens (szakértői-típusú) matematikai programozási modellezési rendszereknél is figyelembe kell venni. A dolgozat második része a matematikai programozási témakör specialitásaiból adódó szempontrendszert tárgyal, mely alapját képezi az ilyen tárgyú modellkezelési rendszerek (model management systems), illetve a matematikai programozási technikákat is felölelő döntéstámogatási rendszerek kidolgozásának. A dolgozat harmadik része rövid irodalmi áttekintést nyújt a matematikai programozási (nagyreszt fejlesztés alatt álló) szakértői rendszerekről. A negyedik rész témája a modellkezelési rendszer indíttatású matematikai programozási szakértői rendszerek felépítésének a problematikája. A dolgozat ötödik része a modellbázis hierarchikus felépítését tárgyalja, megadva a matematikai programozási modellek egy lehetséges osztályozását. A hatodik rész témája a megoldóbázisba tartozó szoftverek rövid, teljességre nem törekvő, elsősorban hazai fejlesztésű, illetve hazánkban hozzáférhető szoftverekre szorítkozó tárgyalása. A dolgozat hetedik része a modellbázis és a megoldóbázis közötti kapcsolatok rövid elemzését tartalmazza. A nyolcadik rész a matematikai programozási tudásreprezentáció kérdéskörét érinti. A kilencedik rész a modellezési nyelvek szempontjából vizsgálja a témakört és a felhasználói interface problematikáját tárgyalja. A tizedik részben a taxonómiai tudás számítógépes megvalósítását ismertetjük. A tizenegyedik rész a rövid összefoglalót tartalmazza.

Mielőtt továbbmennénk szeretnénk néhány alapfogalmat tisztázni.

Modellező rendszer (modeling system) alatt olyan szoftver rendszert értünk, melynek segítségével bizonyos feladatosztályok esetén kényelmesen, a megoldó programcsomagok input-adatformátuma és a programcsomagokban implementált algoritmusok részletes ismerete nélkül lehet a feladatokat megadni és megoldani. A modell megfogalmazása egy deklaratív modellezési nyelv segítségével történik, melynél

a modell megadása szimbolikus és az adatok megadása relációs táblákra támaszkodik. A matematikai programozási modellezési nyelvek általában a matematikai (algebrai) leíráshoz állnak közel. A numerikus modellt a fordítóprogram transzformálja a megoldó programcsomag által megkívánt formátumba (pl. MPS). A fordításból, megoldásból és az eredmények visszajelzéséből álló ciklust egy monitor vezérli. Több megoldó rendelkezésre állása esetén az aktivizálandó programcsomag kiválasztása a felhasználó feladata, akinek erre a célra speciális utasítások állnak a rendelkezésére a modellezési nyelvben.

Modellkezelő rendszer (model management system) alatt olyan szoftver rendszert értünk, amely a modellezési folyamat egészének a támogatását tűzi ki célul, így tartalmaz eszközöket a modellek felépítéséhez, átépítéséhez, tárolásához-visszakereséséhez, megoldásához, a megoldások analíziséhez és az értelmezéséhez, stb. A modelleket mind szimbolikusan, mind numerikusan képes a rendszer kezelni. A modellek reprezentálása történhet modellezési nyelvre építve vagy erre a célra kidolgozott reprezentációs rendszer alapján, ld. Geoffrion (1987). A modell-manipulációs operátorok implementálása vagy egy modell-manipulációs vagy pedig egy interaktív menüvezérelt felhasználói interface-re épül. Technikailag a modell-manipulációs alrendszer általában objektum-orientált. A modellkezelő rendszerek főbb alkotórészei: a modell-manipulációs rendszer, a modellbázis és a megoldó programcsomagokból álló megoldóbázis.

Döntéstámogató rendszer (decision support system) alatt olyan modellkezelő rendszert is tartalmazó szoftver eszközt értünk, amely egy szakterület saját terminológiájában képes az adott, esetleg nem struktúrált vagy csak félig struktúrált problémát intelligensen megközelíteni, megoldani, következtetéseket levonni és ezáltal rávilágítani a probléma természetéből adódó alapvető sajátosságokra. Mivel a döntéstámogató rendszerek nem elsősorban számítástechnikai szakemberek számára készülnek, ezért a felhasználóbarát eszközök (grafika, spreadsheet, stb.), valamint a tudásreprezentációra épülő következtetési és magyarázó intelligens eszközök szerepe lényegesen megnő.

Szakértői rendszer (expert system) alatt olyan programterméket értünk, amelynek segítségével egy szűk szakterületen szimulálni lehet egy szakértő tevékenységét.

Ennek alapján azt lehet mondani, hogy a dolgozat tárgya annak vizsgálata, milyen követelményeket kell a matematikai programozási modellezési és modellkezelési rendszereknek teljesíteni ahhoz, hogy matematikai programozási szakértői rendszereket nyerjünk, amelyek intelligens eszközként szolgálhatnak matematikai programozási feladatok megoldására szakembereknek és nem szakembereknek egyaránt, önállóan vagy más rendszerbe építve.

2. A matematikai programozási szakértői rendszerekkel kapcsolatos néhány szempont

Most a matematikai programozási rendszerek konkrét megvalósításával kapcsolatos néhány szempontot elemezzük, amelyek már működő rendszerek vizsgálatából

adódtak.

2.1 A döntési folyamat és a matematikai programozási rendszer

A matematikai programozási rendszereknél különösen gondosan kell kidolgozni azokat a részeket, amelyek csatlakozási pontot jelentenek a döntéshozatali problémákhoz. A visszacsatolás — mint azt Bíró és Csáki (1988) javasolja — például egy problémaszpecifikus rendszer közvetítésével valósulhat meg, ahol a probléma specifikálása interaktív interjúk segítségével történik.

A modellekre és a megoldó módszerekre vonatkozó tudás reprezentációját úgy kell megválasztani, hogy az a részlegesen specifikált problémák osztályba sorolását támogassa, majd ennek megtörténte után visszajelzést adjon, hogy mely attribútumokra vonatkozó kérdéseket célszerű a felhasználónak feltenni a modell további finomításához, illetve a megfelelő megoldó kiválasztásához.

Ez további kérdéseket vet fel:

2.1.1 A modellek alkalmazhatóságának előfeltételei

Lineáris programozás esetén erre példa az, hogy a modellben a célfüggvény is a feltételi függvények is lineárisak, a döntési változók pedig folytonosak. Az ilyen típusú információk általában a modellek elkülönítő jegyeiként jelentkeznek és az osztályba soroláson keresztül jelennek meg a rendszerekben.

2.1.2 A módszerrel elérhető eredmények minősége

Ezzel kapcsolatban konkrét tények mellett (pl. az eredmény bizonytalansága), procedurális tudás is lehet attribútum, például olyan pointer formájában, amely valamilyen eljárásra mutat. Ez lehet pl. egy modellre vonatkozó érzékenységvizsgálat vagy megoldások létezését kimutató eljárások.

2.1.3 A konzisztencia ellenőrzése

Itt is az egyes egységek attribútumaiként szereplő eljárásokról van szó, amelyek a konzisztencia ellenőrzését végzik.

Természetesen itt nem soroltuk fel az összes lehetőséget.

2.2 A modellezési környezet és a matematikai programozási rendszer

A rendszer kidolgozásakor a teljes modellezési környezetet kell figyelembe venni, amely a modellek közötti kapcsolatok rögzítésén túl a megoldó módszerekkel és az adatforrásokkal való kapcsolatok tisztázását is igényli. A tudásreprezentációt ezért úgy kell megválasztani, hogy a modellezési környezet említett elemeivel és a felhasználói interface-szel együtt kezelhető struktúrákat szolgáltatasson.

Ezzel kapcsolatban a következő fontosabb szempontokat kell figyelembe venni:

2.2.1 A modellbázis és a megoldóbázis közötti kapcsolat

A modellek osztályba sorolását úgy kell végrehajtani, hogy az egyes modellosztályok attribútumaiként lehessen szerepeltetni a megoldókönyvtárral való kapcsolatot.

2.2.2 A modellezési környezet és a matematikai programozási rendszer közötti kapcsolat

Az előzővel analóg követelmény, amely minden nagyobb tematikai egység esetén a kapcsolat újragondolását igényli.

2.2.3 Az adatok hozzárendelése a megoldóbázishoz és a modellbázishoz

Ez a követelmény is hasonló az előzőekhez, de itt a problémát nehezíti az a tény, hogy a különböző megoldó programcsomagok input adatformátuma gyakran nagyon eltérő. A megoldás egy lehetséges módja az, hogy a különböző szintek között transzformációs eljárások biztosítják az átmenetet.

2.2.4 Az alkalmazhatóság határai

Ezzel kapcsolatban attribútumként szerepelhetnek például az aktuális hardver és a megoldó szoftver által szabott méretkorlátok.

2.3 A rendszer rugalmassága

A rendszer rugalmasságán a következő típusú változtatásoknak áttekinthető módon történő, a tudásreprezentációs rendszer által támogatott beépíthetőségét értjük:

2.3.1 A modellbázis kiegészítése új modellekkel

Ez a változás új modellek beépítését jelenti a rendszerbe, ami moduláris, objektumorientált felépítést igényel.

2.3.2 A megoldóbázis kiegészítése új megoldókkal

Ez az előzőhöz teljesen analóg követelmény.

2.3.3 Új vagy módosított attribútumok

Az egyes egységek attribútumainak megváltoztatása szükségessé válhat új modell vagy új megoldó beépítése során.

2.3.4 A modellbázis és a megoldóbázis közötti kapcsolat módosítása

A rendszer használata során, illetve a döntési problémákból származó döntési modellek megoldásakor számos új tapasztalatot nyerünk, amelyeknek a rendszerbe való beépítése annak hatékonyságát, megbízhatóságát növeli.

2.3.5 Adaptáció specifikus döntési problémákra

A rendszert célszerű úgy felépíteni, hogy alkalmas részrendszerek kiválasztásával mód legyen specifikus döntési folyamatokhoz a szükséges modellezési/módszertani tudásbázis alapját előállítani. Ez a redukció azért fontos, mert a tudásbázis használata során, akár lekérdézésről, akár következtetésről van szó, ily módon lényeges gyorsítás érhető el.

2.4 A rendszer egységessége

A matematikai programozási modellek és a módszerek tárgyalásmódja nagyon különböző lehet a történeti kialakulásuk, illetve a támogatott döntési folyamatok

jellege szerint. A modellek nagy része elvileg egységes formai keretek között is tárgyalható, azonban ennek megvalósítása egy matematikai programozási rendszeren belül csak komoly nehézségek leküzdése árán vihető végbe.

A rendszer egységességének alapkövetelménye az egységes tudásreprezentáció.

3. A szakirodalomban publikált rendszerek áttekintése

A matematikai programozási modellek jellemzője, hogy egy vagy több célfüggvény optimális értékét és az ehhez (ezekhez) tartozó döntési változók értékeit kell meghatározni úgy, hogy a döntési változók a korlátozó feltételeket is kielégítsék. A döntéstámogatás segítése szempontjából fontos matematikai programozási modelcsoportba elsősorban azok a modellek vannak besorolva, melyekben a feltételrendszer függvényekre vonatkozó egyenlőségek és egyenlőtlenségek formájában van megfogalmazva. A modellek osztályozása ezután részben a függvények típusa, részben a döntési változók típusa szerint történik (Csáki, Mayer és Szelke, 1988).

A döntéstámogatás segítése szempontjából fontos matematikai programozási modellek a fenti kritériumok alapján a következő négy osztályba sorolhatók:

1. *Lineáris programozás*
2. *Egészértékű (diszkrét) programozás*
3. *Nemlineáris programozás*
4. *Sztokasztikus programozás*

A matematikai programozáson belül, bármely az előbb ismertetett modellosztály esetén lehet tekinteni annak többcélú változatát, azt az esetet, mikor a matematikai programozási feladatnak több célfüggvénye van. Ebben az esetben a főbb alcsoportok a következők:

- I. *Többcélfüggvényes optimalizálás*
- II. *Célprogramozás*
- III. *Hierarchikus programozás*
- IV. *Kompromisszum programozás*

Ezeknek az osztályoknak a pontos meghatározása megtalálható a modellbázisról szóló részben.

A szakirodalomban publikált szakértői alapú rendszerek közül csak azokkal foglalkozunk, amelyek a fentemlített feladatok közül valamelyiket kezeli. Egy gyakorlati szempont volt a válogatásnál, hogy azt a mikrogépes szoftverek körén belül csak az IBM PC XT/AT és az ezzel kompatibilis számítógépekre készült programrendszerekre szűkítettük le, amelyek kellő általánosságban is felhasználható, professzionális rendszerek. Elsősorban azokkal foglalkozunk részletesebben, amelyeknél szerepet kapnak a tudásbázisok és az arra épülő következtetési mechanizmusok.

A második részben a gyakorlatban bevált, standard programokról lesz szó, melyek valamely az előzőekben definiált optimalizálási feladatosztályba tartozó problémát oldanak meg. Egyik esetben sem törekszünk a teljességre, inkább csak képet szeretnénk adni a jelenlegi helyzetről.

3.1 Szakértői alapú rendszerek

3.1.1 EMP

AZ EMP (Expert System for Mathematical Programming) (Schittkowski, 1987) interaktív programrendszer a feltételes matematikai programozási modellek építésére és megoldására szolgáló szakértői rendszer.

A matematikai programozási feladat függvényei lehetnek lineárisak, kvadratikusak, lehetnek adott függvények összegei és maximumai, vagy általános, struktúra nélküli függvények. Így a következő matematikai programozási feladatokat lehet a rendszeren belül megfogalmazni:

- lineáris programozás,
- kvadratikus programozás,
- nemlineáris legkisebb négyzetek problémája,
- nemlineáris vektoroptimalizálás,
- min-max optimalizálás,
- nem-differenciálható optimalizálás és
- általános nemlineáris programozási feladat.

Minden problémánál lehetnek alsó és felső korlátok is, mind a változókra, mind pedig a feltételekre vonatkozóan. Minden optimalizálási probléma megoldásához több megoldó algoritmus programja áll rendelkezésre. Az EMP jelenlegi verziójában a módszerek a következő típusúak:

- szekvenciális kvadratikus programozási módszerek,
- Newton és kvázi-Newton módszerek,
- ellipszoid módszer,
- duál és primál kvadratikus programozási módszerek és
- büntetőfüggvényes módszerek.

Az EMP rendszerben egy programgenerátor is rendelkezésre áll, ami a felhasználó által megadott modellt kapcsolja össze a megfelelő algoritmussal vagy algoritmusokkal.

A modellépítésnél a lineáris és a kvadratikus függvények megadása vektorok és mátrixok megadására egyszerűsödik le, ahol a mátrixok ritkasága, illetve struktúrája is figyelembe vehető. A rendszer a nemlineáris és a nemkvadratikus függvények gradiensét numerikusan számolja, de lehetőség van az analitikus forma megadására is. Ha a különböző matematikai programozási feladatok célfüggvényei vagy feltételi függvényei csak egy indexben térnek el, akkor a függvények megadása lényegesen egyszerűsödik.

A felhasználó a feladat specifikálásához szükséges adatokat a képernyőn keresztül, egy menürendszer segítségével, interaktív módon viheti be. Az általános függvényeket FORTRAN utasítások segítségével lehet megadni és így a függvény nevéhez egy értéket rendelhetünk hozzá. Minden a rendszer által generált probléma bekerül egy adatbázisba, ahonnan bármikor előhívható akár módosítás, akár egyéb célból vagy kitörölhető onnan. Az EMP matematikai programozási szakértői rendszer a megadott probléma megoldására javasol egy megfelelő algoritmust

és generál egy komplett FORTRAN forrásprogramot. A rendszer végrehajtja ezt a programot, majd a kapott eredményeket tárolja egy adatbázisban, így az a későbbiekben is hozzáférhetővé válik. Mivel bármely, a feladatban szereplő függvénynek és változónak lehet nevet adni, ezért a megoldásról problémafüggő eredményt szerkeszthetünk.

A felhasználó kérheti, hogy a rendszer által generált programban változtatásokat hajtson végre, akár saját programegységeinek a programhoz kapcsolásával, akár a generált program egyes szubrutinjainak kicserélésével. A rendszer a feladatok megoldása során szerzett tapasztalatokat összegyűjti, így a működése a tapasztalatok növekedésével jobbá válik. Ha a program a feladatot nem tudja megoldani, akkor egy szabályrendszer által vezérelt hibaanalízis adhat magyarázatot és ennek alapján a hiba kijavítható.

Az EMP programrendszer utasításrendszerét egy menürendszeren keresztül lehet működtetni. Ezekkel az utasításokkal lehet kérni a feladatokhoz tartozó adatok beolvasását, módosítását, törlését, a probléma editálását vagy megoldását, az egyes feladatokhoz tartozó adatbázisok egymásba való transzformációját, a problémáról és a probléma megoldásáról készülő jelentés generálását, a program futásának megállítását, a rendszerről, a matematikai programozási problémákról és a megoldó algoritmusokról információt. Ezenkívül lehetőség van az operációs rendszer parancsainak a bevitelére is.

Az EMP rendszer lehetővé teszi a matematikai programozási eszközök kényelmesebb használatát, így a felhasználó a megoldandó probléma minél pontosabb megfogalmazására koncentrálhat, az adatok minél pontosabb megadására anélkül, hogy a megoldás részleteibe be kellene avatkoznia.

Az EMP rendszernek saját utasításrendszere van, amelynek neve SUSY (Schittkowski, 1987), s amely az IBM PC gépeken kívül használható VAX/VMS és UNIX környezetben is. A SUSY nyelv, amely egyszerűen megtanulható, feltételezi a FORTRAN nyelv ismeretét.

3.1.2 NAXPERT

A NAXPERT szakértői rendszert 1983 óta fejlesztik a szerzők (Schulze és Cryer, 1987). Az alapkiindulásuk az volt, hogy az IBM PC gépeken meglévő mintegy 50 matematikai FORTRAN rutint egy olyan rendszerbe foglalják, amit a matematikában járatlan felhasználó is sikerrel tud alkalmazni a feladatok megoldására, s amely egy matematikában járatos felhasználónak olyan FORTRAN programot vagy FORTRAN program törzset generál, amellyel a felhasználó kísérletezhet, a programot módosíthatja és szükség esetén a programhoz hozzá tudja kapcsolni a saját szubrutinjait.

A NAXPERT rendszer PROLOG nyelven lett implementálva, a tudás kulcsszavakhoz van kapcsolva és a következtetési mechanizmus szabályrendszeren keresztül érvényesül.

A NAXPERT rendszer behívásakor párbeszéd kezdődik a felhasználóval. Ha a felhasználó első alkalommal hívja a rendszert, akkor a használati utasítás jelenik meg először. Ezután egy menürendszer segít kiválasztani a megfelelő menüpontot.

A legáltalánosabban használt menüpont a „konzultáció”. Itt kerül sor a megfelelő rutin kijelölésére. Ehhez kulcsszavakat kell a gépnek megadni. A rendszer kétféle kulcsszó típust ismer:

- lényeges kulcsszó és
- kiegészítő kulcsszó.

A lényeges kulcsszavakhoz vannak kapcsolva mindazok a feltételek, amelyek teljesülése mellett a kiválasztott rutint futtatni lehet. A kiegészítő kulcsszavakhoz tartoznak az addicionális információk.

A NAXPERT rendszer addig nem javasol rutint, amíg minden, a lényeges kulcsszavakhoz tartozó feltétel nem teljesül.

A felhasználó a párbeszéd során a következő esetekben kap segítséget a rendszertől:

- ha a kérdésre nem tud válaszolni, akkor a képernyőn megjelennek a lehetséges válaszok,
- ha nem ismeri pontosan a megfelelő kulcsszót, akkor annak egy részét megadva, a rendszer kiválasztja mindazon kulcsszavakat, amelyek ezt a rész-string-et tartalmazzák.

Ha a felhasználó nincs tisztában egy kulcsszó pontos jelentésével, akkor a help-rendszeren keresztül kérhet pontosabb információt.

A konzultáció a következő két részből áll:

- a megfelelő matematikai részterület kiválasztása és
- a matematikai részterületen a megfelelő rutin kiválasztása.

A NAXPERT rendszer tudásbázisa a következő részekből áll:

- a matematikai részterületek listája (14 részterület),
- a kulcsszavak listája (kb. 160 kulcsszó),
- a szabálybázis (kb. 50 horne-close típusú szabály),
- a megoldó rutinok leírása és
- a rutinokban szereplő paraméterek leírása.

A tudásbázisban szereplő bármely információ könnyen hozzáférhető a felhasználók és azoknak a szakértőknek a számára is, akik azt módosítani akarják.

3.1.3 Egy lineáris programozási szakértői rendszer

Az ebben a részben ismertetendő nemlineáris programozási szakértői rendszer (Chian Kao, 1988) elsősorban a nemlineáris programozási feladatok megoldásánál fellépő egyik nehézséget, a megfelelő megoldó módszer kiválasztását igyekszik megkönnyíteni.

A szakértői rendszer négy részből áll. Ezek a következők:

- felhasználói interface,
- a probléma analízátor,
- a tudásbázis és
- a megoldóbázis.

A probléma analízátor tulajdonképpen egy egyszerű következtető gép, amely a probléma néhány tulajdonsága alapján (változók száma, egyenlőség feltételek

száma, egyenlőtlenség feltételek száma) felépített egyszerű HA AKKOR típusú szabályokból álló szabálybázis segítségével választja ki a megfelelő algoritmust, majd indítja el a megfelelő programot. Ha a kiválasztott algoritmus nem tudja megoldani a feladatot, akkor új algoritmust választ egészen addig, míg meg nem oldja a feladatot vagy amíg el nem fogynak a rendelkezésre álló módszerek.

A szakértői rendszer FORTRAN nyelven íródott, a feladatban szereplő függvényeket is FORTRAN utasításokkal kell megadni. A gradiensek numerikus közelítéseit a rendszer maga számolja a centrális differenciák módszerével. A szakértői rendszer felhasználói interface-e is igen egyszerű, csak a FORTRAN nyelv lehetőségeire épít.

3.1.4 GXMP

A GXMP (Generalized eXperimental Mathematical Programming) (Dolk, 1986) jelenleg lineáris programozási modellezésre alkalmas, azonban a nyelv konstrukciója lehetővé teszi a kiterjesztést a tágabb matematikai programozási modellek, így a nemlineáris programozási modellosztályok irányába is. A modell absztrakciós mechanizmusa lehetővé teszi a tudásbázis alapú felépítést és a különböző felhasználói szintek (döntéshozó, modellépítő, programozó) szerinti struktúrák kialakítását. Viszonylag egyszerű struktúrákat enged meg és az XMP lineáris programozási rendszerre (Marsten, 1981) alapoz. A tudásreprezentáció tudáskeret rendszerű.

3.1.5 LP Formulátor

Az LP Formulátor (Murphy és Stohr, 1986) nagyméretű lineáris programozási feladatokat megcélzó nyelv. Nem algebrai jellegű, hanem a blokkstruktúrák szerinti felépítést támogatja a kisméretű, pl. az időbeli periódusoknak megfelelő, részmodellek igazolása útján. A blokkok egymással való összefüggését hierarchikus hálózati struktúrával reprezentálja, ahol az élek azoknak a soroknak és oszlopoknak felelnek meg, amelyek a részmodelleket összekötik. A lineáris programozási modell felépítése blokkok megadása, a kapcsolatok megadása ezek között, a blokkok finomítása, stb. szekvencia szerint történik. A struktúra felépítésével párhuzamosan, egy szabálybázisú szakértői rendszer segítségével épül fel a feladat mátrixa. A számítógépes implementáció PROLOG nyelven készült. Az adatokkal való kapcsolat a blokkokra vonatkozó adatsémák segítségével történik.

A fentiekben körvonalazott rendszerek a szakértői rendszerek technológiájának három különböző felhasználási módját illusztrálják a számítógépes modellezési környezetekben. Az EMP és a NAXPERT rendszer a problémához legjobban illeszkedő megoldó kiválasztását támogatja, a GXMP rendszer a modellre vonatkozó tudást helyezi a középpontba, míg az LP formulátor a lineáris programozási probléma struktúrált megadását köti össze ilyen technikák segítségével a numerikus megoldáshoz szükséges mátrix generálásával.

A most következő három rendszer gyakorlati feladatokhoz kapcsolódik.

3.1.6 Tartószerkezetek optimális tervezése I.

A tartószerkezetek optimális tervezése témakörben a felmerülő nemlineáris programozási modellek megoldásához a megfelelő megoldó algoritmus kiválasztását tá-

mogatja Hartmann (1986) rendszere. A szabálybázisú rendszer szabályai a különböző problémátípusok és a megoldók egymáshoz rendelését írják le. A rendszer az INSIGHT+ keretrendszerben készült, ld. Level Five Research (1986) és Bokor, Tanyi és Gáspár (1988).

A modellekre vonatkozó tudás elsősorban a modellekben szereplő függvényekre vonatkozó heurisztikus tudás (pl. erősen nemlineáris függvény) és a kiválasztási szabályok az algoritmusokhoz kapcsolódó empirikus tapasztalatokra épülnek.

3.1.7 Tartószerkezetek optimális tervezése II.

A MODLAN rendszer (Papalambros, 1987) szintén a tartószerkezetek optimális méretezése témakör nemlineáris programozási feladatainak modellezési és megoldó problematikájához kapcsolódik. Tartalmazza az előző rendszerhez analóg modell-megoldó hozzárendelést támogató modult és további modulokat, amelyek analizálják a modellt (pl. redundáns feltételek kiszűrése). Mindezen túlmenően az algoritmusokhoz szükséges gradienseket a REDUCE (Rayna, 1987) szimbolikus formulakezelő rendszer felhasználásával kapcsolja az algoritmusokhoz. Ez a probléma tipikus a nemlineáris programozási modellek vonatkozásában: a megoldó algoritmusok nagy része igényli a feladatban szereplő függvények gradienseit, ezek megadása a döntéshozatali környezet esetében nem szabad, hogy szerepeljen a felhasználótól várt információk között. A rendszer szabálybázisú, és PROLOG-ban készült.

3.1.8 Termeléstervezés

A termeléstervezés döntési környezetére vonatkozó modellezési rendszert dolgozott ki Binbasioglu és Jarke (1985). A számunkra lényeges tudásreprezentációs aspektusból nézve a rendszer a tudáskeret reprezentációival dolgozik. A tudáskeretek (frame-ek) egy „is-a” hiererchiába vannak rendezve. A rekeszek (slotok) között vannak „if-needed” típusúak is a származtatott mennyiségek kiszámítására, valamint olyanok, melyek azokra a szimbolikus matematikai összefüggések alakjában megadott feltételekre mutatnak, melyekben szerepel a „slot” által reprezentált objektumnak megfelelő paraméter. Az adatok hozzárendelése relációs táblák segítségével történik.

3.2 Optimalizálási programok

3.2.1 Lineáris programozási rendszerek

Az IBM PC gépeken futó lineáris programozási rendszereken belül csak azokat tekintjük professzionálisnak, amelyek 500 sorosnál nagyobb, általánosan megfogalmazott feladatokat is meg tudnak oldani, és input adatformátumként vagy a 70-es évek elejére ipari standarddá vált MPS-formátumot fogadják el, vagy valamely fejlett modellezési nyelven lehet az inputadatokat megadni.

3.2.1.1 MILP

A MILP programrendszer (Maros, 1988) általános alakban felírt, folytonos lineáris programozási feladatok megoldására szolgál. Használható IBM PC/XT, illetve AT és a velük kompatibilis számítógépeken.

Az optimalizáló algoritmus a felsőkorlátos módosított szimplex módszer alapján működik, multiple pricing technikát használ, és fel van szerelve egy speciális első fázisú eljárással (Maros, 1986). Ezenkívül néhány újszerű algoritmus technikát is kínál. Az eljárás a bázis inverzét szorzat alakban tárolja. Adattárolási és adatkezelési technikája elsősorban nagyméretű, kis kitöltöttségű feladatok hatékony megoldását segíti elő.

A MILP programrendszer a megoldandó lineáris programozási feladatot két különböző formában tudja fogadni: MPS-formátumban és egy saját input formában, ami főleg adatelőkészítő programok által generált feladatok esetén előnyös.

A hazai fejlesztésű MILP a nemzetközi összehasonlításban is előkelő helyen áll. Az European Journal of Operational Research 1989 évi márciusi számában Thiriez közli egy összehasonlító vizsgálat eredményét. Ebben a résztvevő lineáris programozási programcsomagok 40 feladatot oldottak meg. A problémák együttható mátrixának méretei 118×343 és 1101×2981 között változtak. Az összesített értékelésben a MILP végzett a második helyen.

3.2.1.2 XPRESS-LP

A MILP programrendszerrel kapcsolatban említett Thiriez (1989) cikkben a lineáris programozási programcsomagok közül az első helyre az XPRESS-LP programcsomag került. Az értékelő véleménye szerint mikrogépeken ez az általa ismert legjobb minőségű lineáris programozási szoftver.

Az XPRESS-LP programrendszer (Dash Assoc.-Bliworth House, 1988) általános alakban felírt, folytonos és egészértékű lineáris programozási feladatok megoldására szolgál. Használható IBM PC/XT, illetve AT és a velük kompatibilis számítógépeken, valamint nagy gépeken UNIX rendszer alatt.

Az XPRESS-LP a következő 4 modulból áll:

- LP-MODEL a modellépítésre szolgáló egység,
- LP-OPT gyors optimalizáló eljárás folytonos lineáris programozási feladatokra,
- IP-OPT optimalizáló eljárás az egészértékű, lineáris programozási feladatokra és
- LP-REPT az optimális megoldás vizsgálatára szolgáló egység.

Mint már említettük, a Thiriez (1989) cikkben ismertetett összehasonlító vizsgálat alapján az XPRESS-LP programcsomag bizonyult a leghatékonyabbnak, amelynél még külön ki lett emelve az input nyelv hajlékonysága. A további sorrend a következő volt: MILP, XA, LINDO és LPS 687.

3.2.1.3 ZOOM/XMP és GAMS-ZOOM

A ZOOM/XMP (Marsten, 1981) programcsomag 0-1 típusú, vegyes egészértékű lineáris programozási feladatok megoldására alkalmas, amelyeknek a struktúrájára semmilyen kikötés nincs. A megoldó algoritmus korlátozás és szétválasztás típusú, amelyik a folytonos lineáris programozási feladat megoldásából indul. A programrendszer FORTRAN nyelvű, a folytonos lineáris programozási feladatokat az XMP

nevű programrendszer oldja meg. Az XMP először negygépre készült, de van PC-változata is, amelyik az egyik legismertebb lineáris programozási rendszer, nagyon széles alkalmazási és referencia körrel. Komoly előnye a moduláris felépítés, amely alkalmassá teszi lineáris rendszereket kezelő algoritmusok konfigurálását különböző típusú modellekhez.

A GAMS-ZOOM a ZOOM/XMP rendszernek a GAMS (General Algebraic Modeling System) modellezési nyelvhez kapcsolt változata, így a modell megadása a GAMS rendszer keretei között történik.

3.2.1.4 GAMS-MINOS

A nagygépre készült MINOS rendszer PC-re adaptált és GAMS rendszerhez kapcsolt változata. Nemlineáris programozási feladatok megoldására is alkalmas, azonban lineáris programozási oldalról is jelentős alkalmazási sikerekkel bír.

3.2.1.5 LINDO/PC

Az előző két lineáris programozási rendszerrel kapcsolatban említett Thiriez (1989) cikkben a LINDO/PC programcsomag a negyedik helyen szerepel. A LINDO/PC (Linear, Interactive and Discret Optimizer) (Schrage, 1984) programrendszer folytonos lineáris programozási és vegyes egészértékű lineáris programozási feladatok megoldására szolgál. A LINDO/PC az adatokat MPS-formátumban vagy algebrai alakban fogadja el egy inputfile-ről vagy egy spreadsheet file-ről. Azonban az input adatok beolvasásánál indexelési lehetőség nincs. A szerzők kiküszöbölték ezt a hiányosságot a LINGO nevű modellezési nyelvben (Schrage-Cunningham, 1988). A LINDO/PC is lehetővé teszi a közbülső iterációk megőrzését vagy a mátrix invertáló program újraindítását. A program futása közben kiírt hibaüzenet után „on-line help” lehetőséget kínál. A feladat megoldását spreadsheet formátumban lehet megnézni.

3.2.1.6 VINO

A VINO programrendszer (Cunningham és Schage, 1987) nagyméretű, folytonos lineáris programozási és nagyméretű, vegyes egészértékű lineáris programozási feladatok megoldására alkalmas.

A VINO programrendszerénél a folytonos és az egészértékű lineáris programozási feladat optimalizálása össze van kapcsolva azokkal a spreadsheet rendszerekkel, amelyek Lotus file-formátummal dolgoznak (pl. Lotus 1-2-3, Symphony). Létezik a MULTIPLAN verzió is. A cellák típusai: döntési-, feltétel- és konstans-cellák, valamint egy maximálizálandó (minimalizálandó) cella. Ez a felosztás természetes módon megfelel a lineáris programozási feladat változóinak, feltételeinek, illetve paramétereinek azzal a különbséggel, hogy a feltételi összefüggéseket nem kell explicite egyenként megadni, mert a rendszer az esetleges egymásba helyettesítéseket automatikusan elvégzi. Nemnegativitást tesz fel a döntési cellákról és minden ezekről függő celláról.

3.2.1.7 WHAT'S BEST

A WHAT'S BEST rendsze (General Optimization, 1986) a LINDO-nak a Lotus 1-2-3 rendszerrel összeépített változata, így ez is spreadsheet jellegű.

3.2.1.8 SAS/OR SOFTWARE

A SAS/OR Software programrendszer (SAS Institute GMBH, 1988) a „SAS System of Software Tools” nevű rendszer részeként rugalmas, jól használható eszköz lineáris programozási, általános hozzárendelési, legrövidebb út meghatározási, maximális hálózati folyam, minimális költségű hálózati folyam problémák megoldására. Ez utóbbi feladatokat bizonyos típusú korlátozó feltételek mellett is tudja kezelni. A programrendszer segítségével szállítási feladatok is megoldhatók. A hálótervezéshez és a project ütemezéshez a szükséges rajzoló eljárásokat a SAS/GRAPH programcsomag tartalmazza.

3.2.1.9 MICROSOLVE

A MICROSOLVE (Holden-Day Inc., 1988) az alábbiakban részletezendő, alapvetően fontos operációkutató problémák megoldására, IBM PC XT/AT mikroszámítógépen kifejlesztett programrendszer.

A lineáris programozási feladat megoldására szolgáló modulja két részből áll: a modellgenerátorból és a módosított szimplex algoritmust használó optimalizálóból. A feladat megoldása után érzékenységi vizsgálat is kérhető.

A hálózati folyam feladatok megoldására készült modul az általánosított minimál költségű problémákat tudja kezelni, azaz a hálózat mindegyik ágához hozzárendelhető az egység költség, az alsó és a felső határ, valamint a nyereség. Így az ilyen típusú gyakorlati problémák egy nagy része kezelhetővé válik. Ugyanezzel a modullal lehet megoldani a szállítási feladatokat és a hozzárendelési feladatokat. Ez a programegység is egy modellgenerátorból és egy megoldóból áll, amelyeket egy általános menürendszerből lehet hívni.

A MICROSOLVE programrendszer ezenkívül meg tud oldani egészértékű lineáris programozási feladatokat, sorbanállási és szimulációs feladatokat, valamint kezelni tud Markov folyamatokat.

3.2.1.10 MINET

A MINET (Maros, 1988) programrendszer minimális költségű, kapacitáskorlátos (úgynevezett transshipment) hálózati folyam problémák megoldására készült. Ezek a problémák speciális struktúrájú lineáris programozási feladatok, amelyeknél az együttható mátrix minden oszlopában legfeljebb egy +1 és egy -1 érték fordulhat elő párban.

3.2.2 Egészértékű (diszkrét) programozás

Az előző részben ismertetett lineáris programozási rendszerek majd mindegyike rendelkezik olyan modullal, amelyik meg tud oldani vegyes diszkrét vagy diszkrét lineáris programozási feladatot. Ezért ebben a részben más típusú, széles körben alkalmazott diszkrét programozási feladatok megoldására szolgáló programrendszereket ismertetünk, pl. Szelke (1988).

3.2.2.1 TRAVEL

A TRAVEL programcsomag (Dept. of Combinatorics and Optimization, the University of Waterloo, 1988) az úgynevezett utazó ügynök probléma megoldására szolgál. A rendszer menüvezérelt és ezen keresztül a felhasználó maga dönti el, hogy melyik algoritmussal kívánja a feladatot megoldani. A programrendszer algoritmus animációt alkalmaz, amelynek segítségével futás közben nyomon lehet követni a megoldó algoritmus működését. Ezzel kapcsolatos a programrendszernek az a hasznos tulajdonsága, hogy lehetőség van a megoldó kritikus pontokon való megállítására.

A programrendszer elsődlegesen az Euklideszről elnevezett utazó ügynök feladat megoldására készült, ami azt jelenti, hogy ebben az esetben az utazás során érintett városok közül bármely kettő távolsága euklideszi. Azonban a rendszer lehetővé teszi egy komplett, szimmetrikus, távolság mátrix explicit specifikációját a felhasználó számára, és ezzel megnyitja az utat a nem-euklideszi, szimmetrikus problémák kezelésére is.

3.2.2.2 TRASAL

A TRASAL program (Naddor és Eliezer, 1988) az utazó ügynök feladat megoldására szolgál. Módszerét tekintve „branch-and-bound” technikát alkalmaz. A program egyike annak a 12 operációkutatási programnak, melyet a Hopkins Egyetem E. Naddor-ék készítettek.

3.2.2.3 Erőforrás korlátos hálótervező modul

Az erőforrás korlátos hálótervező modul (Bíró, 1988) a korlátozott mennyiségben rendelkezésre álló erőforrások optimális allokálását végzi el. Ez a hálótervezési tevékenység egyik fontos részfeladata, ugyanis egy könnyen előállítható, lehetséges allokálást választva, az így kapott hálóterv időtartama általában jóval hosszabb az erőforrás korlátok nélküli kritikus időtartamánál.

A modul segítségével előállítható egy olyan erőforrás korlátos hálóterv, amelynek az időtartama a lehető legrövidebb. E feladat megoldásának a nehézségi foka lényegesen meghaladja a hálótervezés többi fázisát és kézzel vagy nem kellően hatékony számítógépes eljárással még kis számú tevékenység esetén is legtöbbször reménytelen.

A modul magja egy hatékony keresési eljárás, amelynek során a hálóterv időtartama lényegesen csökkenhet, közelítve az erőforrás korlátok nélküli kritikus időtartamhoz, amelyet csupán a tevékenységek közötti logikai kapcsolatok határoznak meg. A keresési eljárás során a képernyőn nyomon követhető az időtartam csökkenése és a keresés interaktívan megszakítható, ha az eredmény kielégítő.

A keresési eljárás automatikus befejeződését okozhatja az erőforrás korlátok nélküli kritikus időtartam esetleges elérése, vagy a megvizsgált teljes hálótervek számának rögzített korlátja.

3.2.3 Nemlineáris programozás

3.2.3.1 EMP

A szakértői alapú rendszerek között a 3.1.1 pontban szerepel.

3.2.3.2 NAXPERT

A szakértői alapú rendszerek között a 3.1.2 pontban szerepel.

3.2.3.3 MINOS és GAMS/MINOS

A MINOS (Murtagh és Saunders, 1987) a legnevesebb nemlineáris programozási programcsomag, amelyik nagyon jó nemzetközi referenciákkal rendelkezik. A MINOS nagygépre készült, de a GAMS-hoz kapcsolt változata minden gépen fut. A GAMS/MINOS egy modellezési rendszer része, ezért használata programozási tudás nélkül is lehetséges (nem kell a függvényszámoló rutint FORTRAN nyelven megírni, a lineáris programozási adatokat MPS-formátumban elkészíteni, stb.).

A matematikai programozási rendszerekben a MINOS használata nagyon elterjedt. Ezt mutatja, hogy 1980 óta kb. 200 különböző intézetben vagy üzemben installálták a rendszert. A MINOS népszerűségének az oka, hogy nemcsak kis-, hanem nagyméretű lineáris programozási, feltétel nélküli optimalizálási, lineáris feltételrendszerű nemlineáris programozási, és általános nemlineáris programozási feladatokat is meg tud oldani. A feladatok megoldásához a simplex módszert Dantzig (1947 és 1963), a kvázi-Newton-módszert Davidon (1959), a redukált gradiens módszert Wolfe (1962) és a vetített Lagrange multiplikátor módszert Robinson (1972), Rosen és Kreuser (1972) használja.

A programrendszer nyelve FORTRAN és ebből következően paraméter vezérelt. Kb. 80 paraméter értéke változtatható. A rendszer szolgáltatásai közé tartozik a numerikus gradiens kiszámítása. A programrendszer újabb változata interaktív, azonban így is csak matematikai programozásban jártas szakemberek tudják használni.

3.2.3.4 NLPSOLVER

Az NLPSOLVER (Powell, 1988) programcsomag általános nemlineáris programozási feladatokat old meg és az optimalizáláshoz a Powell-módszert használja. Matematikai programozásban jártas felhasználók számára készült, akik gyakorlatok a programozásban és ezt a modult képesek beépíteni a saját programjukba. A programcsomag FORTRAN nyelvű. A feladatok megadását matematikai jelölésmódban kéri. Ennek megkönnyítésére az indexelés lehetővé van téve.

3.2.3.5 Egy nemlineáris programozási szakértői rendszer

A szakértői alapú rendszerek között a 3.1.3 pontban szerepel.

3.2.4 Sztochasztikus programozás

A sztochasztikus programozási feladatok numerikus szempontból a nehéz feladatok kategóriájába tartoznak. Megoldásuk nagyméretű lineáris programozási feladatok vagy nehéz nemlineáris programozási feladatok megoldását jelenti. Ez utóbbi esetben a nehézség nem feltétlenül a méretből adódik, hanem abból, hogy a feltételi

függvények és a gradiensek kiszámítása nagyon számításigényes és csak viszonylag pontatlanul lehetséges. Az elkészült sztochasztikus programozási programrendszerek ezért gyakran nagygyépen futnak. Az alábbiakban ismertetendő programrendszereknek van IBM/PC-n futó változatuk is.

3.2.4.1 QDECOM

A QDECOM (Ruszczyński, 1986) az úgynevezett „fixed recourse” kétlépcsős feladatok megoldására készült, ahol a második lépcső célfüggvényében, együttathatómátrixában és a jobb oldalán egyaránt szerepelhetnek valószínűségi változók. A valószínűségi változók eloszlása véges diszkrét, a programrendszer FORTRAN nyelvére és a PC-n futó változata az XMP-t is használja. Az alkalmazott algoritmus Ruszczyński regularizált dekompozíciós módszere (Ruszczyński, 1986), aki a programrendszer kifejlesztője is.

3.2.4.2 TWOSTAGE

A TWOSTAGE (Strazicky, 1987) programcsomag kétlépcsős sztochasztikus programozási probléma megoldására szolgál, ahol a második lépcsőben szereplő jobb oldali vektor, illetve együttatható mátrix elemei lehetnek diszkrét eloszlású valószínűségi változók. Az alkalmazott algoritmus Strazicky kompaktbázis módszere. A programcsomag FORTRAN nyelvére.

3.2.4.3 PCSP

A PCSP (Szántai, 1985, 1988) programcsomag lineáris feltételrendszerű, lineáris célfüggvényű feladatokat old meg, ahol a feltételrendszer még egy megbízhatósági jellegű feltételt is tartalmaz. A valószínűségi feltételben lineáris feltételrendszer áll, amelynek jobb oldala véletlen vektor és amelyről feltételezzük, hogy együttes eloszlásuk normális. A megoldó algoritmus Veinott-metszősík módszerének továbbfejlesztése. A valószínűségi függvény értékek és a gradiensek kiszámítása a Monte Carlo szimulációs technika segítségével történik (Szántai, 1986). A programrendszer FORTRAN nyelvére.

3.2.4.4 PROCON

A PROCON (Mayer, 1988) programcsomag olyan feladatokat tud megoldani, ahol a lineáris feltételrendszert egy megbízhatósági jellegű feltétel egészíti még ki. A valószínűségi feltételben is lineáris feltételrendszer áll, amelynek a jobb oldala véletlen vektor. Erről feltételezzük, hogy az együttes eloszlása normális. A megoldó algoritmus redukált gradiens típusú. Két változata van. Az egyiknél a lineáris feltételek kezelése a MILP programrendszerrel történik, a másikon pedig az XMP-vel. A programrendszer FORTRAN nyelven íródott.

4. A matematikai programozási szakértői rendszerek felépítéséről

A mai igényeknek megfelelő matematikai programozási szakértői rendszereknek az alábbiakban részletezendő követelményeknek kell megfelelnie. A pontosításra

azért van szükség, mert a szakértői rendszereknek nincs általánosan elfogadott definíciója és ezért sokkal több szoftvert kiáltanak ki annak, mint amennyi valójában van. Ebben a cikkben szakértői rendszer alatt olyan ismeretalapú programrendszert értünk, amelyik valamely, pontosan körülhatárolt (általában szűk) tárgyterületen belül, gyakorlati szinten használható fel feladatmegoldásra. Az itt megfogalmazott követelményrendszert Geoffrion (1987) és Maturana (1987) cikkeinek felhasználásával állítottuk össze. Az ebben szereplő követelmények egymástól nem függetlenek, hanem bizonyos részekben egyik a másikat tartalmazza.

1. *A matematikai programozási modellek számítógépes reprezentációja legyen kényelmesen kezelhető a végfelhasználó számára, feleljen meg a kellő matematikai szigorúságnak, pontosságnak, valamint tegye lehetővé a megoldó programok hatékony használatát.*

Ez a szempont komplex módon tartalmaz több másikat is, mint az a későbbiekben kiderül.

2. *A modellbázis és a megoldóbázis legyen független.*

A matematikai programozási szakértői rendszerek modellbázisában nagyon sok különböző modell lehet, amelynek mindegyikéhez több megoldó rutint lehet rendelni. Ezek a rutinok azonban közel sem egyforma hatékonysággal működnek az egyes feladat típusokon, sőt vannak olyanok, amelyek nem is tudják megoldani az adott feladatot. Ennek számos oka lehet, amelyek egy része előre nem látható, ezért a módszer kiválasztását nem érdemes előre lerögzíteni és a modellt, valamint az optimalizáló egységet egybeépíteni. Ez azzal az előnnyel is jár, hogy az optimalizálási módszerek fejlődésével az újabb programváltozatokat minden különösebb erőfeszítés nélkül beépíthetjük a rendszerbe. Így egy úgynevezett „nyitott” megoldókönyvtárat kapunk.

3. *A matematikai programozási szakértői rendszer modellezési nyelve legyen elég általános ahhoz, hogy a tárgykörbe tartozó problémák könnyen modellezhetőek legyenek.*

A matematikai programozási szakértői rendszerek hasznossága és hatékonysága nagymértékben függ a rendszer modellezési nyelvétől. Ennek tulajdonítható, hogy eddig is már számos modellezési nyelv született, pl. AMPL, CAMPS, EMP, GAMS, GINO, GXMP, HEQS, IFPS, LINDO, LP-MODELL, MLD, PAM, PLATOFORM, UIMPS, WHAT'S, BEST!, stb. A modellezési nyelvekről részletesebb tájékoztatást ad Csáky, Meyer és Szelke (1988).

A modellezési nyelv akkor megfelelő, ha elég hatékony a modellezésben, és könnyű megtanulni. Ezeket a szempontokat azért nehéz kielégíteni, mert egymásnak ellentmondóak. Gondoljunk például az indexelési lehetőségekre, ami a modell matematikai leírását nagyon megkönnyíti, ezzel szemben a modellezési nyelv megtanulását megnehezíti. Egy másik példa a beépített függvények száma és típusai.

4. A modellbázis kezelése legyen megfelelő szinten megoldva.

A matematikai programozási szakértői rendszereknek általában nagyszámú modell kezelésére kell alkalmasnak lennie, azaz képesnek kell lennie ezeket a modelleket tárolni, újból aktivizálni és az esetleg több felhasználó részéről történő módosításokat figyelembe venni. Ez azt jelenti, hogy a modellbázissal hasonló műveleteket kell végezni, mint egy adatbázissal:

- új modelleket kell beépíteni a rendszerbe,
- modelleket kell a rendszerből kitörölni és
- modelleket kell a képernyőn megjeleníteni.

Ezenkívül speciális, csak a modellekre vonatkozó műveletek is vannak. Ezek a következők:

- modellekből adódó feladatok megoldása,
- modellek összekapcsolása,
- modellek összeépítése,
- modellek dekomponálása és
- modellek vizsgálata.

Vegyük észre, hogy még az adatbázisokkal megegyező műveletek esetén is lényegi különbség van a modellbázis és az adatbázis műveletei között, ugyanis míg a modellbázis esetén a műveletek szimbolikus, addig adatbázis esetén numerikus és karakter feldolgozás-orientáltak.

5. Az adatkezelés legyen megfelelő szintű.

Az előző pontból kiderül, hogy a modellkezelés és az adatkezelés az operációkat tekintve különböző, de az egységes elveken nyugvó struktúrák lehetnek hasonlóak. Az ebben a pontban szereplő szempontot megerősíti az a tény, hogy a matematikai programozási feladatok megoldásához általában negyötömegű input adatra van szükség.

A jól működő rendszereknél az a törekvés figyelhető meg, hogy a szükséges input adatokat a legegyszerűbb formában kelljen megadni és ebből a rendszer állítsa elő a modell megoldásához szükséges adatbázist. Ezekben a rendszerekben az adatkezelést relációs adatbáziskezelő végzi.

6. A modellbázis és a megoldóbázis közötti kapcsolat biztosítsa az optimalizálási feladatok automatikus megoldását.

A matematikai programozási rendszerek legfontosabb része az optimalizálást végző megoldóbázis. Ismert azonban, hogy még a nagyon jó optimalizáló programok is, amelyek kihasználják a megoldandó feladatok specialitásait, csak egy szűk modellosztályon belül működnek kielégítően. Ezért minden modell esetén felvetődik a kérdés, hogy a megoldóbázis melyik rutinja segítségével próbáljuk megoldani a feladatot, s ha azzal nem sikerül, akkor melyik legyen a következő, és mikor mondjuk azt, hogy a rendszer nem tudja megoldani az adott feladatot?

Ennek alapján úgy tűnik, hogy érdemes a modellbázis és a megoldóbázis között egy univerzális interface-t kialakítani, amelynek segítségével nem okoz gondot az

egyik algoritmusról a másikra való áttérés. A lineáris programozási rendszereknél ez már megvalósult, az MPS-formátum éppen ezt a szerepet tölti be.

A modellbázis és a megoldóbázis közötti kapcsolat nem csak az adatátvitelen keresztül valósul meg, hanem egy szabályrendszeren keresztül is, amelyben a feltételek teljesülése vagy nem teljesülése alapján választja ki a rendszer a megoldó rutint.

A szabályrendszer mellett egy procedurális rész is tartozhat az interface-hez, amelynek segítségével lehetőség van a feladatok mélyebb, struktúrális vizsgálatára. Példa lehet az, mikor a döntési változókat (a döntéshozók számára lényeges információkat tartalmazó, pl. kontroll) a többi változótól elkülönítve tekintjük és csak az előbbiek függvényében vizsgáljuk a feladat linearitását. Így sok esetben nemlineáris programozási feladat helyett lineáris programozási feladatot kell megoldani, ami lényegesen könnyebb.

Csak az előbbiek figyelembevételével képzelhető el olyan interface a modellbázis és a megoldóbázis között, amelyik hatékonyan tudja biztosítani az optimalizáló program automatikus kiválasztását.

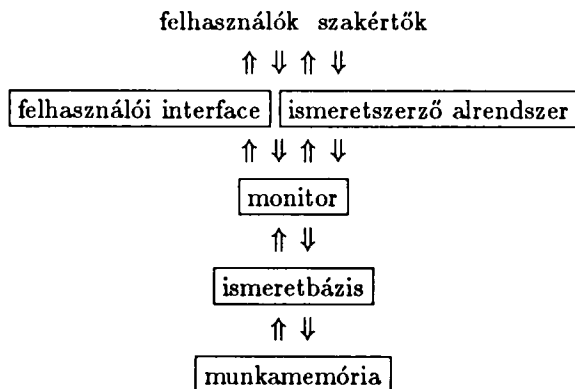
7. A felhasználói interface feleljen meg a mai követelményeknek.

A felhasználói interface-nek építenie kell a ma használatos legmodernebb technikákra (pl. Windows ablakkezelő rendszer, grafika, stb.), ezzel is megkönnyítve a rendszer megtanulását és használatát. Ezenkívül támogatnia kell a tudásbázis építését és a feltöltését.

8. Legyen a rendszer moduláris felépítésű.

Az előbbiekben vázolt rendszer kiépítése, módosítása és továbbfejlesztése csak moduláris rendszerben képzelhető el, ahol az egyes egységek között különböző erősségű és minőségű kapcsolat lehet. Ezeket a kapcsolatokat nagymértékben meghatározza a cél és a rendszer felhasználási köre.

A matematikai programozási szakértői rendszerek vázlatos felépítése látható az 1. ábrán.



1. ábra

A rendszerben a monitor vezérli és ellenőrzi az egyes egységek működését. Idetartozik az úgynevezett következtető gép és a magyarázó-indokló alrendszer is. Mint látható, külön egység biztosítja a szakértővel való kapcsolatot.

Az ismeretbázis a modellkezelő rendszeren kívül a modellbázist és a megoldóbázist is tartalmazza.

Az eddigiekből is látható, hogy IBM PC gépeken nem valósítható meg olyan matematikai programozási szakértői rendszer, amely minden a gyakorlatban fellépő, érdekes matematikai programozási feladatot hatékonyan meg tudna oldani. Ezért a reális cél csak az lehet, hogy egy szakértői rendszer vázát adjuk meg, amelyből a konkrét alkalmazási feladatok ismeretében lehet jól használható eszközt kifejleszteni.

Az elmondottakat két példán szemléltetjük. Ezek a példák mutatják azt, hogy a szakértői rendszer céljának változása milyen mélyreható változásokat eredményezhet a programrendszerben.

Az első esetben tegyük fel, hogy magas szintű, stratégiai döntéseket szeretnének hozni egy vállalatnál. Ez azt jelenti, hogy a fellépő matematikai programozási feladatokat csak néhányszor kell megoldani, és sem a megoldás pontossága, sem pedig a gyorsasága nem lesz döntő szempont. Döntő lesz viszont az, hogy mennyire könnyű egy új modellt beépíteni a rendszerbe, egy régit módosítani vagy kitörölni onnan, a modelleket összekapcsolni, aggregálni vagy esetleg dekomponálni. Ebből látszik, hogy ilyen esetben különösen gondosan kell a modellbázis kezelését megvalósítani.

A második esetben tegyük fel, hogy a szakértői rendszert ugyanennél a vállalatnál szeretnék felhasználni, de a termelésben, az operatív döntések előkészítésénél. Ebben az esetben előre tudni lehet, hogy milyen típusú, milyen méretű feladatokat kell megoldani és a feladatok megoldásához mennyi idő áll rendelkezésre. Ezért itt arra kell törekedni, hogy a megoldást minél gyorsabban, minél pontosabban kapjuk meg a fellépő néhány speciális feladatra. Ehhez elsősorban kiváló megoldó rutinokra és adatkezelő rendszerre van szükség.

Ezek a példákon keresztül is látszik, hogy a szakértői rendszerek meghatározásánál nem véletlenül szerepel az a kitétel, hogy valamely szűk tárgyterületen belül alkalmazható csak feladatmegoldásra. Ehelyett talán helyesebb azt mondani, hogy a szakértői rendszer fogalmához hozzátartozik az a pontos cél is, amelyre készült.

5. A modellbázis

A matematikai programozási szakértői rendszerek fő célja az alkalmazások során felvetődő matematikai programozási feladatok minél pontosabb osztályozása, majd ennek alapján minél hatékonyabb megoldása. Ez azt jelenti, hogy először egy matematikai, hierarchikus függés szerinti osztályozás történik, amelyben szerepet kap az algoritmus kiválasztást támogató tudás is.

Egy matematikai programozási szakértői rendszer az alábbi, általános formában megadott feladatok osztályozására és ezek közül bizonyos előre kiválasztott

osztályokba tartozó feladatok megoldására szolgál.

$$\begin{aligned} \min f(x) \\ g_i(x) &\geq 0, \quad i = 1, \dots, m, \\ h_j(x) &= 0, \quad j = 1, \dots, p, \\ x_a &\leq x \leq x_f, \\ x &\in R^n, \quad \text{vagy } x \in E^n, \end{aligned}$$

ahol

- n a változók száma,
- m az egyenlőtlenség feltételek száma,
- p az egyenlőség feltételek száma,
- R^n az n -dimenziós Euklideszi tér,
- E az egész számok vagy azoknak egy részhalmaza,
- x_a és x_f a változókra vonatkozó alsó és felső korlátok.

Ezenkívül azt is feltételezzük, hogy a feladatban szereplő függvények minden változójukban folytonosak. A rendszernek tudnia kell kezelni azokat a feladatokat is, ahol bizonyos feltételekre mindkét oldalról van korlátozás, de ezeket is a fenti formába lehet transzformálni. A matematikai programozási modellek, mint az előzőekben már láttuk, a következő négy nagy csoportba sorolhatók:

1. *Lineáris programozás*
2. *Nemlineáris programozás*
3. *Egészértékű (diszkrét) programozás*
4. *Sztokasztikus programozás*

Ezek egy csoporton belül a modellek további tulajdonságait, valamint a megoldó algoritmusokra vonatkozó tudást figyelembe véve újabb részosztályokat jelölhetünk ki. Mivel az osztályozás függ a rendszer céljától is, ezért itt csak a legalapvetőbb és legjellemzőbb osztályokat ismertetjük.

Az osztályokon belül, az egyes részosztályok között egységes interface-t tételezünk fel, megkönnyítve így az egyikből a másikba való átmenetet. Az osztályozásnál tekintettel voltunk arra a tényre is, hogy minden újabb osztály vagy részosztály bevezetése, a már létező megoldó programot (programokat) figyelembe véve, olyan transzformációs programok írását teszik szükségessé, amelyek az egyes osztályokhoz tartozó egységes ismeretbázist transzformálják át a másinak megfelelő alakba, illetve az adatokat a kiválasztott megoldó által megkövetelt formára hozzák. Mivel egy szakértői rendszerben a modelleknek legalább három reprezentációja van (egy a végfelhasználó számára, egy a matematikai programozási rendszer modellbázisában és egy a kiválasztott algoritmushoz), így látható, hogy mennyi nehézséget jelent újabb osztályok bevezetése, illetve mennyire nem hatékony egy túlságosan nagy rendszer.

A matematikai programozási rendszerek részletes osztályozására taxonómiai program készült (Vermes, 1989), amelynek ismertetése a 10. pontban található.

5.1 Lineáris programozás

A lineáris programozási modellosztályon belül három részosztályt érdemes külön tekinteni. Az első kettő annyira alapvető, hogy minden matematikai programozási rendszerben helye van. Ezek a következők:

5.1.1 Folytonos lineáris programozási modell

Folytonos lineáris programozási modellnek nevezzük azt a matematikai programozási modellt, amelyben a célfüggvény is és a feltételi függvények is lineárisak, valamint a döntési változók folytonosak. A matematikai programozási feladatosztályon belül ehhez kapcsolódnak a legnagyobb alkalmazási sikerek, amelyeknek egyrészt az a titka, hogy számos gyakorlati probléma ilyen alakban fogalmazható meg, másrészt az, hogy nagyméretű feladatok is megoldhatók numerikusan és a megoldások minőségére, a megoldások érzékenység vizsgálatára, a modell analízisére több lehetőség közül választhat a felhasználó.

Ezekről részletesebben olvashatunk, pl. Murty (1976) és Greenberg (1983) könyveiben.

5.1.2 Egészértékű lineáris programozási modell

Egészértékű (diszkrét) lineáris programozási modellnek nevezzük azt a lineáris programozási modellt, amelyben a döntési változók között olyan is szerepel, amelyik csak egész értékeket vehet fel. Ezen belül megkülönböztetünk tiszta diszkrét (valamennyi változó egészértékű) és vegyes diszkrét feladatokat. A tiszta diszkrét feladatokon belül pedig úgynevezett 0 – 1 feladatokat, melyeknél a diszkrét változók csak 0 vagy 1 értékeket vehetnek fel.

A matematikai programozási rendszerekben általában 0 – 1 típusú tiszta vagy vegyes, lineáris, diszkrét feladatokat megoldó rutinokat találhatunk, ugyanis a többi lineáris, diszkrét programozási feladat ilyenre transzformálható.

Ezekről a feladatokról részletesebben olvashatunk, pl. Nemhauser és Wolsey (1989) könyvében.

5.1.3 Speciális lineáris programozási modell

Nagyon sok területen, széleskörűen alkalmaznak olyan modelleket, amelyek speciális struktúrájú lineáris programozási feladatok megoldására vezetnek. Ezeknek a feladatoknak a megoldására, amennyiben a rendszer céljával ez összhangban van, érdemes külön rutint (rutinokat) beállítani.

A speciális lineáris programozási modelleken belül az alábbiakat különböztetjük meg:

- szállítási feladat,
- hálózati folyam feladat,
- hozzárendelési feladat,
- blokkdiagonális, pl. Dantzig–Wolfe struktúrájú feladat,
- diszjunktív programozási feladat és
- dekomponálható struktúrájú feladat.

5.2 Nemlineáris programozás

Nemlineáris programozási modellnek nevezzük azt a matematikai programozási modellt, amelyben a célfüggvény és a feltételi függvények közül legalább egy nem lineáris és a döntési változók folytonosak. A modellcsoportba tartozó modellek száma igen nagy, pl. Bazaraa és Shetty (1979) és Luenberger (1976) könyvei.

Általában egy nemlineáris programozási feladat megoldása komoly nehézséget jelent. Ezért célszerű a megoldhatóság szempontjából külön osztályba sorolni a méretek szerint, a különböző nehézségű feladatokat. Így három osztályt lehet kialakítani:

Az elsőbe a kisméretű feladatok tartoznak, amelyeknél a változók száma nem nagyobb, mint 5, a feltételek száma pedig, mint 10.

A másodikba a középméretűek, amelyeknél a változók száma nem nagyobb, mint 50, a feltételek száma pedig, mint 100.

A harmadikba a nagyméretű feladatok sorolhatók, azaz azok, amelyek az előző osztályokba nem tartoznak bele.

A matematikai programozási feladatok ilyen felosztása a számítógépes tapasztalatokon nyugszik, ezért a mérethatárok csak közelítő jelleggel érvényesek. Ezen belül is prioritást kell biztosítani a változók számának, amely általában jobban jellemzi a probléma nehézségi fokát, mint a feltételek száma. Ezt az osztályozást a tapasztalatok birtokában tovább lehet finomítani minden, a továbbiakban bevezetendő részosztályon belül is. Ez lehet a „tanulás” egyik formája.

A nemlineáris programozási feladatok egy másik lehetséges osztályozása, a struktúra szerinti osztályozás, amelynél a feltételrendszer típusa szerint a következő három részosztályt különböztetjük meg:

5.2.1 Feltétel nélküli feladatok

A feltétel nélküli feladatok körén belül érdemes különválasztani az egy és a többváltozós feladatokat.

5.2.2 Lineáris feltételrendszerű feladatok

A lineáris feltételrendszerű feladatok vagy a célfüggvény típusa szerint oszthatók további alcsoportokra vagy a lineáris feltételrendszer struktúrája szerint.

A célfüggvény típusa szerinti csoportosítás a következő:

- kvadratikus,
- hiperbólikus,
- konvex,
- entrópia és
- konkáv.

Természetesen ezeken kívül (vagy ezek helyett) is képezhetünk még osztályokat, ha a szakértői rendszer célja úgy kívánja.

A lineáris feltételek struktúrája szerinti osztályozás hasonló, mint a lineáris programozásnál.

5.2.3 Nemlineáris feltételrendszerű feladatok

A nemlineáris feltételrendszerű feladatok között külön csoportba tartoznak a következők:

- konvex programozási feladatok,
- geometriai programozási feladatok,
- komplementaritási feladatok,
- szeparábilis feladatok és
- nemdifferenciálható feladatok.

Ezekon kívül természetesen még nagyon sok csoport képezhető, pl. bármely kettő kombinációja.

5.3 Egészértékű (diszkrét) programozás

Egészértékű (diszkrét) programozási modellnek nevezzük azt a matematikai programozási modellt, amelyben a döntési változók között szerepel olyan, mely csak egész értékeket vehet fel. Ezen belül megkülönböztetünk tiszta diszkrét (valamennyi változó egészértékű) és vegyes diszkrét feladatokat.

A diszkrét modellekre vonatkozóan bővebb információkat találhatunk, pl. Murty (1976), Nemhauser és Wolsey (1989) könyveiben.

Mivel egy matematikai programozási szakértői rendszernél tekintettel kell lenni a bevezetendő részosztályoknál a felhasználói interface-re, ezért a lineáris diszkrét feladatokat, a hasonló interface miatt, a lineáris programozásnál tárgyaltuk.

Másrészt általános, közép méretű, diszkrét programozási feladat megoldása is sok esetben reménytelen, ezért ebben a részben csak struktúrált feladatokkal foglalkozunk.

Ezek a következők:

- hátizsák feladat,
- darabolási feladatok (egy-dimenziós darabolás, két-dimenziós darabolás, Guillotin-vágás, két-dimenziós darabolás, nem Guillotin-vágás),
- ütemezési feladatok (sorrendezési modellek, erőforrások ütemezése, órarend modellek, menetrend modellek) és
- struktúrált halmazokkal kapcsolatos feladatok.

5.4 Sztochasztikus programozás

Sztochasztikus programozási modellnek nevezzük azt a nemlineáris programozási modellt, amelyben a célfüggvény és a feltételi függvények közül legalább egy valószínűségi változót is tartalmaz. A döntési problémák szempontjából legfontosabb sztochasztikus programozási modellek a lineáris programozási modellből származnak olyan módon, hogy annak bizonyos együtthatóit valószínűségi változóként kezeljük. Az erre a döntési helyzetre kidolgozott matematikai programozási modelleket sztochasztikus lineáris programozási feladatoknak nevezzük. Ezek a modellek a matematikai programozás szemszögéből általában vagy nagyméretű lineáris programozási, vagy nemlineáris programozási, vagy vegyes-diszkrét lineáris feladatok a valószínűségelosztástól és a modell megfogalmazásától függően. A modellek a numerikus megoldás szempontjából a nehéz feladatok közé tartoznak.

A részletekre vonatkozóan bővebb információ található pl. Kall (1976) könyvében.

A két legfontosabb modelltípus a következő:

- valószínűséggel korlátozott (megbízhatósági) modell és
- többlepcsős modell.

A valószínűséggel korlátozott modellek alosztályai az egyedi valószínűségi feltételekre vonatkozó, illetve az együttes eloszlásra épülő modellek.

5.5 Többcélú feladatok

A matematikai programozáson belül, bármely az előbb ismertetett modellosztály esetén lehet tekinteni annak többcélú változatát, azaz azt az esetet, mikor a matematikai programozási feladatnak több célfüggvénye van. Ebben az esetben a főbb alcsoportok a következők:

5.5.1 Többcélűfüggvényes optimalizálás

A matematikai programozáson belül többcélűfüggvényes optimalizálásról akkor beszélünk, ha az úgynevezett efficiens pontok meghatározása a cél, illetve ezek halmazának bizonyos tulajdonságú pontjai.

5.5.2 Célprogramozás

A matematikai programozáson belül célprogramozásról akkor beszélünk, ha a különböző célfüggvények előírt értékeit szeretnénk megkapni (goal) és ezért olyan megoldást keresünk, melyre az eltérés ettől a céltől minimális.

5.5.3 Hierarchikus programozás

A matematikai programozáson belül hierarchikus programozásról akkor beszélünk, ha előre meg van adva a célfüggvények egy prioritási sorrendje és először a legfontosabbat véve optimalizálunk, majd ezt optimális értéken rögzítve, előírjuk, mint feltételt, vesszük a következő célfüggvényt, stb.

5.5.4 Kompromisszum programozás

A matematikai programozáson belül kompromisszum programozásról akkor beszélünk, ha az úgynevezett ideális ponthoz valamilyen metrikában a legközelebbi pont meghatározása a cél.

A többcélű optimalizálással kapcsolatban részletesebb információk találhatók Zeleny (1982) könyvében.

6. A megoldóbázis

Ebben a részben a matematikai programozási szakértői rendszerek megoldóbázisáról lesz szó, melyben a rendszerhez tartozó modelltípusok megoldását célzó megoldó programrendszerek találhatók. Ezt a következő szempontok figyelembevételével kell kialakítani:

- minden a modelbázisba tartozó alapeladat esetén, lehetőség szerint több megoldó program álljon a rendelkezésre, ugyanis csak így lehet biztosítani a rendszer megbízható működését,

- a modelltípusok megoldásánál nem algoritmusokra, hanem programokra kell javaslatot tenni, mert egy algoritmus hatékonysága nagymértékben függ az implementációtól,

- a rendszerbe kerülő programok jó nemzetközi referenciákkal rendelkezzenek,

- a programok lehetőség szerint minél egyszerűbben legyenek beépíthetőek a matematikai programozási szakértői rendszerekbe,

- nagyméretű feladatok megoldására csak a pontos cél ismeretében lehet érdemi javaslatot tenni.

6.1 Lineáris programozás

6.1.1 Folytonos lineáris programozás

A folytonos lineáris programozási feladatok megoldásához az XMP, a MINOS és a MILP programrendszerek használatát javasoljuk.

6.1.2 Egészértékű lineáris programozás

Az egészértékű lineáris programozási feladatok megoldásához a ZOOM és a MILP 88 programrendszerek használatát javasoljuk. A ZOOM programrendszer 0 - 1 típusú, vegyes egészértékű feladatok megoldására alkalmas. A folytonos lineáris programozási feladatok megoldására az XMP rendszert használja. A MILP 88 egészértékű lineáris programozási feladatokat old meg.

A programok jellemzőiből látszik, hogy a MILP 88 kisebb méretű, általános egészértékű lineáris programozási feladatokat kezel, míg a ZOOM rendszer inkább azokat, amelyekben több, de csak 0 - 1 típusú egész változó van.

6.1.3 Speciális struktúrájú lineáris programozás

Az előzőekből kiderült, hogy nagyon sokféle, speciális struktúrájú feladat megoldása merülhet fel egy szakértői rendszerben, amelyeknél ráadásul a megoldásra fordítható idő a rendszer céljától is függ. Elképzelhető viszont az is, hogy a szakértői rendszer célja olyan, hogy a megoldandó feladatnál nem szükséges kihasználni a speciális struktúrát, pl. a feladat mérete kicsi. Ezért nem érdemes előre megmondani azt, hogy a különböző struktúrájú feladatokhoz milyen programot rendeljünk hozzá.

Ebben a részben csak egy hazai termékre, a hálózati folyamatok kezelésére alkalmas MINET programsomagokra teszünk javaslatot.

6.2 Nemlineáris programozás

6.2.1 Feltétel nélküli feladatok

A feltétel nélküli nemlineáris programozási feladatok speciális esetei a feltételes feladatoknak, ezért kis- és közép méretben a feltétel nélküli problémák megoldhatók a feltételes feladatok megoldására szolgáló programokkal.

Amiért érdemes külön osztályba sorolni ezeket a feladatokat, az az, hogy nagyméretű feladatok esetén, a feladat pontos megfogalmazásának az ismeretében, nagyon hatékony algoritmusokat lehet kidolgozni, másrészt, ha a rendszer szolgáltatásai között szerepel a programgenerálás is, akkor a rendszer lényegesen rugalmasabb lesz, ha nemcsak többdimenziós, hanem egydimenziós feltétel nélküli optimalizáló rutinok is rendelkezésre állnak.

6.2.2 Lineáris feltételrendszerű feladatok

A nemlineáris programozási döntési problémák közül jelenleg a lineáris feltételrendszerű feladatok a legfontosabbak. Ez a feladatosztály ugyanis nagyon közel áll a legnagyobb alkalmazási sikerekkel büszkélkedő lineáris programozáshoz, így annak sok előnyével rendelkezik (pl. a feltételrendszer standardizált adat inputja, amelyik a lineáris programozásával teljesen megegyezik) és nem nagyon érvényesülnek azok az okok, amelyek miatt az általánosabb nemlineáris programozási modellek nem elég széleskörűen használatosak.

Ehhez a feladatosztályhoz egy nagyon jó nemzetközi referenciákkal rendelkező programrendszert fejlesztettek ki, amelynek neve MINOS (Murtagh és Saunders, 1978 és 1982).

A lineáris feltételrendszerű problémákon belül talán a legfontosabbak a kvadrátikus programozási feladatok. Ennél a modellosztállynál csak három hazai fejlesztésű programcsomag megemlítésére szorítkozunk.

A BEALE nevű program a Beale-módszer implementációja, amelyik a szimplex módszerhez hasonlóan működik (Bernau, 1979).

A BEALEUP program a Beale-módszer módosításának az implementációja, amellyel lehetővé válik a változókra vonatkozó egyedi alsó és felső korlátok kezelése is (Bernau, 1979). Ezek a programok nagyon jó nemzetközi referenciákkal rendelkeznek (Schittkowski, 1987).

A NEQP program a konvex kvadrátikus programozási feladatokra kifejlesztett Wolfe-módszer implementációja. Ennél az implementációnál a feladatban szereplő változókra nincsenek előjelkorlátok (Soós, 1986).

6.2.3 Nemlineáris feltételrendszerű feladatok

A nemlineáris feltételrendszerű feladatok közül a szakértői rendszer céljának a pontosítása nélkül a középmező, általános nemlineáris programozási feladatok a legfontosabbak. Ezek megoldására javasoljuk az alábbi, széles körben elterjedt szoftvereket:

- az előző részben ismertetett MINOS programrendszer kiterjesztését nemlineáris feltételrendszerű feladatokra és az
- NLPSOLVER programcsomagot, amely a Powell-módszer FORTRAN nyelvű implementációja.

A programcsomag nem végfelhasználók, hanem gyakorlott programozók számára készült, imputként matematikai jelölésmódban leírt feladatot is elfogad és a probléma leírására az indexes jelölést is lehetővé teszi.

A hazai programcsomagok közül középmező, általános nemlineáris programozási feladatok megoldására szolgálnak a következők:

- az l_1 SQP programcsomag, amely szekvenciális, kvadratikus programozási algoritmussal dolgozik, a változókra megadható egyedi alsó és felső korlátokat elkülönítve kezeli. A teszteredmények alapján, a nemzetközi értékítélet szerint a program az NLPSOLVER-rel megegyező minőségű (Bernau, 1987) és

- az RQPLNP program, amely szintén szekvenciális, kvadratikus programozási algoritmussal dolgozik (Soós, 1986).

A lineáris komplementaritási problémát oldja meg a LEMKE program, amelyik a Lemke-módszer egy módosítására épül (Bernau, 1983).

A többi nemlineáris feltételrendszerű, nemlineáris programozási feladatra a konkrét cél ismeretében érdemes megfelelő minőségű szoftvereket kiválasztani.

6.3 Egészértékű programozás

Az egészértékű feladatok megoldásához is, ha azok nem a lineáris egészértékű modellek családjába tartoznak (ezekről korábban részletesebben szoltunk), csak célszoftvereket érdemes figyelembe venni. Ezek közül itt kettőt említünk meg:

- a TRAVEL programrendszer az utazó ügynök feladatot oldja meg és
- az intézetünkben fejlesztett, erőforrás korlátos hálótervező modul a korlátozott mennyiségben rendelkezésre álló erőforrások optimális alokálását végzi (Bíró, 1989).

6.4 Sztochasztikus programozás

A sztochasztikus programozási feladatok megoldására nem áll rendelkezésre széleskörűen elterjedt, általános célú szoftver, mint pl. a MINOS a lineáris és a nemlineáris programozás területén. Ezért csak három szoftvert ismertetünk, melyből kettő az intézetünkben készült a IIASA-val kötött kutatási szerződés, illetve a Zürichi Egyetem Operációkutatási Intézetével kötött együttműködés keretében.

A PCSP és a PROCON programcsomag olyan valószínűséggel korlátozott (megbízhatósági jellegű) sztochasztikus programozási feladatok megoldására készült, melyeknél a valószínűségi feltételben csak a jobb oldali vektor sztochasztikus és a jobb oldal együttes eloszlása normális. Az implementált megoldó algoritmusok Veinott metszősík módszerére, illetve a redukált gradiens módszerre épülnek (Szántai, 1988 és Mayer, 1988). A PROCON-ban a lineáris feltételek kezelése a MILP, illetve az XMP segítségével történik.

A TWOSTAGE programcsomag olyan kétlépcsős feladatok megoldását célozza, ahol a második lépcsőben a jobb oldal és a mátrix sztochasztikus és az eloszlás véges diszkrét. Az implementált algoritmus kompaktbázis technika az ekvivalens nagyméretű, lineáris programozási feladatra (Strazicky, 1987).

6.5 Egyéb programok

A megoldóbázishoz tartoznak az előbb említetteken kívül azok a programok is, amelyek a megoldást előkészítik, segítik, hibaanalízist végeznek, stb.

7. A modellkezelési rendszer

A matematikai programozási szakértői rendszereknél alapkövetelmény, hogy a modellbázissal kapcsolatban is lehetővé váljanak egyrészt olyan operációk, amelyek az adatbázisoknál is előfordulnak (pl. tárolás, visszakeresés, betöltés, listázás, katalogizálás, archiválás, stb.), másrészt specifikusan csak itt előforduló tevékenységek, pl. megoldó kiválasztása, a megoldás érzékenységvizsgálata, a modell átépítése, modellvariánsok ajánlása a felhasználó számára, az eredmények ellenőrzése, stb.

A modellbázisnál az adatbázisszerű operációk is új tartalommal telnek meg és sokkal bonyolultabbak lesznek. Vegyük például a visszakeresést és a betöltést. A modell azonosításánál nem elég az adatbázisoknál elfogadott kulcsokkal dolgozni, a visszakeresés a taxonómiai rendszer üzemeltetésén keresztül történik. A modell betöltése a szokásostól annyiban tér el, hogy a modellt leíró adatstruktúra sokkal összetettebb, hiszen egy frame-t kell kitölteni. Ez a numerikus adatok mellett tartalmazhat szimbolikus adatokat (pl. függvények, ahol a betöltés után automatikusan szimbolikus deriválásra is sor kerülhet), szabályok ellenőrzése történik meg és a hiányzó adatokat ennek alapján kezdi keresni a rendszer (pl. aktivizálja a dialog management rendszert vagy az adatbázisokban kezd keresni).

Ha a matematikai programozási szakértői rendszer egy döntéstámogató vagy egy másik szakértői rendszer részét képezi, akkor biztos, hogy nem a matematikai programozásban jártas szakembereknek készült, ezért a modellhez a megoldó program kiválasztását érdemes automatikussá tenni, azaz a matematikai programozási rendszernek „fekete dobozként” kell működnie úgy, hogy elég nagy biztonsággal, pontossággal és gyorsasággal oldja meg a feladatot.

A jelenlegi matematikai programozási rendszereket megvizsgálva megállapíthatjuk, hogy ez a cél még messze van. Ennek a fő okai a következők:

- nagyon sokféle matematikai programozási modell van, ezért ezeknek egy rendszeren belüli kezelése nehéz feladat, különösen ha nemcsak osztályozásról, de megoldásról is szó van,
- minden modelltípushoz több, esetenként nagyszámú megoldó program közül választhatunk,
- a jelenlegi rendszerekben a megoldó kiválasztása a modellhez elsősorban formai jellemzők alapján történik (változók száma, feltételek száma, stb.),
- a rendelkezésre álló, jóminőségű szoftverek többsége általános célú, csak közepes méretű feladatok megoldására alkalmas és
- a rendszerek intelligenciája nem mindig felel meg a mai lehetőségeknek (pl. modellezési nyelv, tanulási képesség, stb.).

A továbbiakban megvizsgáljuk, hogy milyen részeken keresztül valósulhat meg a modellbázis és a megoldóbázis közötti kapcsolat.

7.1 Adatbázis

A matematikai programozási szakértői rendszerekben a modellbázis és az adatbázis jöllehet különböző kezelést igényel, mégis célszerűnek tűnik a numerikus adatokat kezelő részt és a taxonómiai rendszert összehangolni. Vegyünk például egy

lineárisan korlátozott nemlineáris programozási modellt. Az első lépésben elkezdődik a modell azonosítása, azaz a taxonómiai fában elindulnak a gyökértől felfelé. (Ez a későbbiekben részletesen ismertetésre kerül.) Az első kapcsolódási pont az adatbázishoz az, mikor megtudjuk, hogy a célfüggvény is lineáris, vagy hogy a feltételek hálózatiak. Mindenesetre abban a pillanatban, amikor kiderül, hogy általános lineáris feltételekről van szó, akkor hozzárendelhető a modellhez a numerikus adatbázisnak megfelelő része. Az öröklődés miatt a taxonómiai fa minden további csomópontján öröklődik ez a kapcsolat, azaz a hozzárendelés a numerikus adatbázis lineáris feltételeket tartalmazó részéhez a többi alárendelt modell számára is meg van oldva. Így a kapcsolat mindig azon a szinten jön létre, ahol annak először van értelme. Ezeken a pontokon szükség lehet a felhasználóval történő kommunikációra a hiányzó adatok megszerzése céljából. Az adatkezelést tehát a taxonómiai rendszer hajtja végre, ezért annak többi részével egységet alkot. Úgy gondoljuk, hogy érdemes hasonló elvek alapján felépíteni mind a két egységet. Az adatkezelésnek ez a módja a döntéstámogató rendszereknél is elterjedőben van.

A taxonómiai rendszerben gondoskodni kell arról, hogy a numerikus adatbázis a kiválasztott megoldónak megfelelő alakra transzformálódjon. Megjegyezzük, hogy az input adatfile általában különbözik a modell adatbázisától, ezért először ezt a transzformációt kell végrehajtani.

7.2 Szabályrendszer

A matematikai programozási szakértői rendszerekben a különböző modellosztályok fastruktúrában helyezkednek el. A taxonómiai osztályozásnál a cél az, hogy egy konkrét modell helyét megtaláljuk ebben a fastruktúrában. A taxonómiai osztályozás a modellek jellemzői alapján, szabályok segítségével történik. Az egyes modellosztályokhoz megoldó programok vannak rendelve és a modellek, valamint a megoldók tulajdonságai alapján, szintén szabályrendszeren keresztül vannak összekapcsolva. Így látható, hogy a modell és a megoldó programok összekapcsolása két lépésben történik és szabályrendszereken keresztül valósul meg.

A rendszer annál hatékonyabban fog működni, minél pontosabban fel vannak derítve egyrészt a modell és a megoldók tulajdonságai, másrészt minél hatékonyabb a szabályok kezelése. Ezért újabban a szabályrendszerek is struktúráltak, azaz a szabályok a feldolgozás szempontjainak megfelelően vannak csoportosítva. Így nem szükséges minden szabályt, minden esetben egyenként megvizsgálni. Ennek különösen nagy rendszerek esetén van jelentősége. Esetünkben a szabályok struktúrálására természetes lehetőségként adódik a taxonómiai fa szerkezetének a figyelembe vétele.

7.3 Modellvizsgálat

A jelenleg működő matematikai programozási rendszerekben a modellek osztályozása, valamint a modellek és a megoldó programok összekapcsolása elsősorban formai jegyek alapján történik, pl. változók száma, feltételek száma, stb.

Újabban egyre inkább követelmény, hogy legyen lehetőség a feladatban szereplő függvények szimbolikus vizsgálatára és pl. egyszerűsítésére. A lineáris programozási feladatok kezelése lényegesen hatékonyabb, mint a többi matematikai programozási

feladaté és mint az előző részekből kiderült, így sokkal nagyobb méretű feladatok oldhatók meg, ezért pl. a függvénykiértékelésnek lehet a célja a függvények lineáritásának a megállapítása.

Mivel nem remélhető, hogy a matematikai programozásra vezető gyakorlati problémák, vagy lineáris programozási feladatok, vagy pedig általános rutinokkal megoldható közép méretű feladatok, ezért az egyik legfontosabb kutatási iránynak a modellek és a megoldó algoritmusok mélyreható vizsgálata tűnik.

Elsősorban a modellek szimbolikus vizsgáltában kellene előbbrelépni. Szükség lenne ugyanis a modell megadása után, többek között eldönteni azt, hogy a feladatban szereplő függvények folytonosak-e, egyszer vagy kétszer differenciálhatóak-e, lineárisak vagy esetleg konvexek, stb.

Ehhez nyújt pl. segítséget az, ha a matematikai programozási feladatban szereplő függvények úgynevezett faktorábilis formában vannak megadva. Ezt a fogalmat McCormick (1983) vezette be, mert a tapasztalatok azt mutatták, hogy a lineáris programozási feladatokon kívül, még közepes méretű feladatok esetén sem garantálható a megoldás. A vizsgálatok szerint az egyik fő ok az, hogy a többváltozós függvények számítógépes reprezentációja nem megfelelő. Ennek következtében mind a függvényértékek, mind a gradiensek és a Hesse mátrixok pontbeli értékeinek kiszámítása, mind pedig a feladatok struktúrális tulajdonságainak a megállapítása vagy nehézkes vagy nem végezhető el.

Ezeknek az okoknak a kiküszöbölésére vezetett be McCormick a faktorábilis függvények fogalmát, amelyeknek a definíciója a következő:

A többváltozós függvényt faktorábilisnek nevezzük, ha megadható véges számú, egyváltozós, elemi függvény, amelyekből az összeadás, a szorzás, a rekúzió és az egyváltozós függvények felhasználásával, véges számú lépésben előállítható a függvény.

Egyszerűen belátható, hogy az ilyen formában megadott függvények értékének, valamint a gradiens és a Hesse mátrix értékének a kiszámítása lényegesen könnyebbé válik, mert a függvény megadása és kiszámítása összhangban van.

7.4 Tanulás

Nemlineáris programozási feladat esetén, még közép méretű problémákat tekintve sem garantálható a megoldás, ezért a rendszer működése szempontjából lényeges a „tanulás”, azaz a rendszernek az a képessége, hogy a problémamegoldás során szerzett tapasztalatokat a továbbiakban hasznosítsa.

A „tanulási folyamatnak” vannak egyszerűbb és bonyolultabb változatai. Az egyik legegyszerűbb talán az, mikor minden megoldóhoz hozzárendelünk egy 0 és 100 közötti mérőszámot, amelyik teljesen szubjektív alapon a program „jóságát” mutatja. Ha a program egy feladatot megoldott, akkor ez a mérőszám egy előre megállapított értékkel nő, ellenkező esetben csökken. Ehhez természetesen párosulhat annak a figyelése is, hogy egy modell, illetve egy megoldó hány feladat megoldásánál kapott szerepet.

A matematikai programozási rendszerekben jelenleg csak ilyen egyszerű módon történik a tapasztalatok feldolgozása. Ennek az egyik fő oka az, hogy a rendszerek

nem képesek komolyabb modell vizsgálatára. Az elkövetkezendő években ezen a területen is nagy előrelépés várható.

További lehetőség a tanulásra a hibaanalízis elvégzése. Ehhez pontosan kell ismerni a modellek és a megoldó algoritmusok, illetve a megoldó programok szerkezetét. Ezenkívül megfelelő hatékonyságú programoknak kell rendelkezésre állni.

Egy újabb lehetőség az úgynevezett algoritmus animáció, amelyik az algoritmus egyes részeinek vizuális megjelenítésével próbál újabb információkra szert tenni, illetve ezeket a továbbiakban hasznosítani. Erről a felhasználói interface-ről szóló részben bővebben lesz szó.

8. Matematikai programozási rendszerek ismeretbázisa

A matematikai programozási rendszerek ismeretbázisáról bővebb információ található Csáki, Mayer és Szelke (1988) dolgozatában. Az előzőekből kiderült, hogy egy matematikai programozási rendszer ismeretbázisa csak struktúrált objektum lehet. Itt a tudáskeret alapú ismeretábrázolásról lesz szó. A tudáskeret struktúrált objektum, amelyhez az attribútumok (rekeszek) lehetséges értékeinek tág skálája tartozhat.

A tudáskeret-rendszer a tudáskeretek egy hierarchikus rendszere, ahol egy adott tudáskeretnek a hierarchiában alárendelt tudáskeret az összes olyan attribútumát örökli, melyek nincsenek saját (own) rekeszként deklarálva. Az alárendelt tudáskeretben a rekesz értéke megváltoztatható, illetve általában megengedett az értéktartomány szűkítése is. Így ha a megelőző tudáskeretben valamely attribútum értékei egy halmazból származnak, az alárendelt tudáskeretben a halmaz részhalmaza is előírható az értékre vonatkozóan.

Mivel a tudáskeret rendszer információ-visszakeresési rendszerként is működik, ezért a rendszerhez hozzátartozik az a cél, amiért a matematikai programozási rendszert létrehozták, azaz a rendszer fő funkciója is. Matematikai programozási szakértői rendszer esetén ezt a fentiekben körvonalaztuk. Ez a cél határozza meg a tudáskeret-rendszer pontos struktúráját és a vezérlés is ennek megfelelően van megtervezve (pl. a tudáskeret-rendszer felhasználása is).

A tudáskeret-rendszerben az egyes rekeszek kitöltésénél, vagy értékének meghatározásánál szükség lehet különböző eljárások végrehajtására (pl. a rekeszek típusának automatikus ellenőrzése). Ezeket az eljárásokat, amelyek bizonyos feltételek teljesülése esetén aktivizálódnak, nevezik a rekeszekben levő démonoknak.

A matematikai programozási rendszerek ismeret ábrázolására az egyik lehetőség tudáskeret típusú rendszer építése. Esetünkben ugyanis a rendszert arra kívánjuk felhasználni, hogy egyrészt parciálisan specifikált modellek esetén megtaláljuk a hierarchiában legmélyebben található olyan tudáskeretet, melynek attribútum értékei konzisztensek a tudásunkkal, azaz asztálybasorolást hajtsunk végre, másrészt ennek alapján ki tudjuk választani a megoldóprogramot (programokat).

Az osztályozás úgy történik, hogy egy adott helyről indulva, a hierarchiában lefelé haladva valamely algoritmus alapján összehasonlítjuk az illető tudáskeret reke-

szeit a modell eddig ismert attribútumaival. Ha eljutottunk a legmélyebben fekvő rekeszhez, akkor ennek vagy az utódainak még specifikálatlan attribútumai modellezési szinten segítenek a modell további specifikálásában. (Utódon mindenütt közvetlen utódot értünk.)

Az osztályba sorolást nagymértékben elősegíti, ha a tudáskeretek egyes rekeszeiben szabályok vannak. Igen hasznos ilyen szempontból elegendő feltételek beépítése arra vonatkozóan, hogy egy modell, amelyik a megelőző tudáskerettel konzisztens, mikor konzisztens az utód-tudáskerettel is.

Mivel a tudáskeretek az öröklődés tekintetében fastruktúrát alkotnak, ezért ez a tudáskeret-rendszer felhasználható a szabályok struktúrálására is, így gyorsabb következtetési eljárások dolgozhatók ki.

Megjegyezzük, hogy a szakértői rendszereknél alapkövetelmény, hogy az ismeretbázis és a monitor külön egység legyen, ezért az egész rendszer vezérlése, az osztályozási algoritmus, a különböző következtetési vagy egyéb eljárások beindítása az ismeretbázison kívül történik.

Az ismeretbázis egyes tudáskereteiben a következő rekeszeket találjuk:

(1) Taxonómiai rekesz

A taxonómiai rekeszek a tudáskeretek helyét specifikálják a tudáskeret-rendszerben, a megelőző-tudáskeret és az utód tudáskeretek azonosítóinak felsorolása útján. Az utód tudáskeretek felsorolási sorrendje kifejezi azt, hogy mennyire gyakran használatos modellről van szó. A felsorolás a döntési gyakorlatban jobban elterjedt módszerekkel kezdődik.

(2) Osztályozó rekesz

Az osztályozó rekeszekben a tudáskeret rekeszeire vonatkozó feltételekből — mint elemi tényekből — konjunkcióval és diszjunkcióval, valamint tetszőleges zárójelezéssel képzett logikai kifejezés áll, a konjunkció tagjaiban a tudáskeret rekeszeire vonatkozó valamely feltételekkel. A kifejezés elegendő feltételt szolgáltat arra vonatkozóan, hogy a tudáskeret megelőző tudáskeretével konzisztens döntési probléma ezzel a tudáskerettel is konzisztens-e, azaz a tudáskeret által képviselt osztályba besorolható-e. Így, ha az osztályba sorolás folyamatában elértünk egy olyan állapotot, melynél a kifejezés True értéket vesz fel, akkor a probléma ebbe az osztályba besorolható. Valójában az osztályozás az inkonzisztens utódok kizárásával fog működni, ezért a logikai kifejezést úgy kell felépíteni, hogy a kiértékelés sorrendje megfelelően a prioritási sorrendnek. Magasabb prioritást azok a feltételek kapnak, melyek nemteljesülése várhatóan előbb derül ki az osztályozás során.

(3) Modell rekesz

Ezekben a rekeszekben találhatók a modelleket jellemző adatok és a modellek számítógépes reprezentációja (pl. méretek, célfüggvény és feltételi függvények, stb.), valamint a modellek tulajdonságai (a feladatban szerep-

lő függvények linearitása, konvexitása, stb.). Ez a rekesz tehát elsősorban a modellkezelés funkcióinak megfelelően épül fel.

(4) Adat rekesz

Az adatrekeszekben találhatók azok a visszakeresési és tárolási módszerek, melyek hozzárendelést biztosítanak az adatbázis azon részeihez, amelyek az aktuális modell megoldásához szükségesek.

(5) Szabály rekesz

Ezek a szabályok a tudáskeret egyes rekeszei között fennálló esetleges logikai összefüggéseket tartalmazzák. Alkalmazásuk során a monitor mindig valamely eddig ismeretlen rekesz értékére következtet. A szabályok alkalmazhatóságának a vizsgálatát a monitor automatikusan végzi.

(6) Eljárás rekesz

Ezekben a rekeszekben találhatók azok az eljárások, amelyek a megoldó algoritmusokat elindítják, különböző, a megoldást előkészítő vagy segítő eljárások (pl. adatelőkészítő, adatellenőrző programok), transzformációkat végrehajtó programok (pl. modell transzformációk, adatbázis transzformációk, stb.).

Az ismeretbázishoz természetesen hozzátartoznak a modellbázis, a numerikus adatbázis és a megoldóbázis is, amelyek a tudáskeretek megfelelő rekeszein keresztül kapcsolódnak a rendszerhez.

9. A felhasználói interface

A matematikai programozási szakértői rendszereknek két felhasználási módja lehet. A matematikai programozásban jártas szakemberek számára érdemes biztosítani a közvetlen hozzáférés lehetőségét, így a rendszer lényegesen hatékonyabban működtethető. Ha viszont a matematikai programozásban kevésbé jártas szakembereknek kívánunk minél kényelmesebben használható, könnyebben megtanulható rendszert bocsátani a rendelkezésére, akkor érdemes megpróbálni egy olyan változatot kialakítani, amelynél a matematikai programozási szakértői rendszer egy másik szakértői vagy döntéstámogatói rendszerbe van építve. Ezért ebben a részben az egyik legfontosabb feladat olyan modellezési nyelv kiválasztása, amely egyrészt lehetővé teszi a minél könnyebb és minél pontosabb modell megadását a matematikai programozási szakértői rendszer számára, másrészt ennek a nyelvnek a struktúrája biztosítja azt, hogy könnyen összekapcsolható legyen pl. egy döntéstámogatási problémakörre épített rendszerrel.

Elképzelhető modellezési nyelv nélküli, menüvezérelt interface is. Ilyenek általában a döntéstámogató rendszereknél találhatók, amelyeknél így elkerülhető a modellezési nyelv megtanulása.

9.1 A matematikai programozási rendszer modellezési nyelve

A matematikai programozási nyelvekről részletesebb tájékoztatást ad Csáki, Mayer és Szelke (1988).

A hagyományos modellezésnél az első lépés az volt, hogy a modellező a felhasználóval együtt megalkotta a probléma egy szimbolikus modelljét. Ez a modell általános volt abban az értelemben, hogy a benne szereplő matematikai szimbólumok egy egész problémaosztályt írtak le. Ezután következett a modell numerikus megadása, azaz a megoldáshoz szükséges adatok megadása, mégpedig olyan formában, ami megfelelt a feladat megoldásához kiválasztott programsomag input előírásainak.

A számítógép szerepének a növekedése először a lineáris programozás esetén volt megfigyelhető. A lineáris programozási megoldó programsomagok input adatformátuma a 70-es évek eleje óta az MPS-formátum. Nagyméretű modellek esetén az adatok kézi megadása már gyakorlatilag megoldhatatlannak bizonyult, és megjelentek a mátrix generátorok. A mátrix generátorok a modellezési nyelvek előfutárainak is tekinthetők, amelyek még csak a lineáris programozási feladat mátrixának a megadására szolgáltak, ugyanis az ebben a nyelvben írt programokat lefuttatva eredményként az MPS-formátumú adatokat kaptuk. A leglényegesebb különbség egy mátrix generátor és egy modellezési nyelv között az, hogy a generátor nyelv procedurális, a modellezési nyelv pedig deklaratív. Ez azt jelenti, hogy a generátor nyelvekben a mátrixokat értékadó-jellegű utasításokkal kell feltölteni, vannak ciklusutasítások, ugró-utasítások, stb., amelyek hasonlónak teszik egy programnyelvhez, pl. FORTRAN, de ezek specializálva vannak mátrixok, illetve lineáris programozási feladatok esetére. A modellezési nyelvek deklaratív jellege azt jelenti, hogy lehetővé válik a feladatok leírása a numerikus adatokkal együtt anélkül, hogy bármilyen procedurális-nyelvi lépést tartalmaznának, így megtanulásuk nem feltételez számítógépes programozási ismereteket.

A mátrix generátorok fő hiányossága az volt, hogy a modell módosítása egy számítógépes program módosítását tette szükségessé. Mivel az utóbbi évtizedekben a lineáris programozási modellek felhasználási köre egyre bővült, ezért ezen a területen jelentek meg először valódi modellezési nyelvek. Ez azt jelenti, hogy a modellezési nyelvek segítségével nemcsak egyszerűen és könnyen megtanulható módon lehetett a modellt szimbolikusan leírni, ami a modell ellenőrzésénél és módosításánál lényeges könnyebbséget jelentett, hanem a modell struktúrájához jól illeszkedő módon az adatokat is meg lehetett adni.

A modellezési nyelvek területén azóta intenzív kutatások folynak. Számos új, elsősorban lineáris programozási modellezési nyelv született:

- AMPL (A Mathematical Programming Language), Fourer, Gay és Kernighan (1987),
- LINDO (Linear, Interactiv and Discret Optimizer), Schrage (1984),
- LPL (Linear Programming Language), Hürlimann és Kohlas (1988),
- LPMODEL (Linear Programming Modeling), Katz, Risman és Rodeh (1980),

- CAMPS (Computer Assisted Mathematical Programming System) Lucas és Mitra (1985) és
- PAM (A Practitioner's Approach to Modeling), Welch (1987);

általános célú modellezési nyelvek:

- GAMS (General Algebraic Modeling System), Bisshop és Meeraus (1982),
- GINO (General Interactive Optimizer), Liebman, Lasdon, Schrage és Waren (1986),
- LINGO Cunningham és Schrage (fejlesztés alatt),
- MLD (Modeling Language and Database), Burger (1982) és
- PLATOFORM (Planning Tool written in Dataform), Palmer (1984);

spreadsheet rendszerű modellezési nyelvek:

- VINO Cunningham és Schrage (1987),
- WHAT'S BEST! General Optimization (1986) és
- IFPS/OPTIMUM Roy, Lasdon és Lordeman (1986);

szakértői rendszerekkel kapcsolatos modellezési nyelvek:

- EMP (Expert System for Mathematical Programming), Schittkowski (1987),
- GXMP (Generalized Experimental Mathematical Programming), Dolk (1986) és
- LP FORMULATOR Murphy és Stohr (1986).

A modellezési nyelv kiválasztásánál szem előtt kell tartani, hogy elég általános modell típusok is leírhatóak legyenek, ezért nem jöhetnek számításba a csak lineáris programozásra készült nyelvek, valamint a spreadsheet rendszerű modellezési nyelvek sem, hiszen a spreadsheet által szolgáltatott modellezési keret erősen behatárolja a lehetőségeket, ha különböző típusú modelleket akarunk kezelni. Ez csak igen erőszakoltan és áttekinthetetlenül valósítható meg.

Az általános célú nyelveket tekintve a GINO a LINDO-ra épül, da a LINDO az indexelési lehetőség hiánya miatt csak kisméretű feladatok leírására alkalmas. A GINO-nak azóta elkészült az újabb változata, amelyik tartalmazza az indexelést.

A LINGO fejlesztés alatt áll, jelenleg csak lineáris és vegyes diszkrét feladatok megoldására alkalmas.

A PLATOFORM nagygépes rendszer.

A szakértői rendszerek technikáját alkalmazó modellezési nyelvek közül az EMP-ben található SUSY a FORTRAN nyelvhez kötődik és elsősorban a heterogén megoldó készletből a megfelelő program kiválasztását és a paraméterek értékeinek helyes beállítását támogatja. A GXMP és az LP FORMULATOR jelenleg lineáris programozási feladatok modellezésére alkalmas.

Mivel a GAMS (General Algebraic Modeling System) modellezési nyelv lehetővé teszi lineáris, nemlineáris és egészértékű feladatok modellezését is, valamint az indexelési lehetőség miatt nagyméretű feladatok kezelésére is alkalmas, ezért modellezési nyelvnek ezt javasoljuk. Ezt a választást támasztja alá az is, hogy az előzőekben javasolt, jó nemzetközi referenciákkal rendelkező programcsomagok közül a MINOS, az XMP és a ZOOM már össze van kapcsolva a GAMS rendszerrel,

és a legkülönbözőbb gépekre implementálva van, ezért a matematikai programozási rendszerben egy egységes interface alapját képezheti.

9.2 A GAMS modellezési nyelv

A GAMS (General Algebraic Modeling System) modellezési nyelvet a Világ Bank fejlesztési projektjeinek a stratégiai tervezésének a támogatására dolgozták ki. Ez azt jelenti, hogy a legkülönbözőbb területeken felmerülő döntési problémák modellezésére alkalmas. Mivel stratégiai tervezésről van szó, ezért a modellek számítógépes reprezentációjának könnyen áttekinthetőnek és érthetőnek kell lennie, az eredményeket a tervezésben résztvevők számára indoklással kell világossá tenni, lehetővé kell tenni más modellekkel való összekapcsolását, a modellek gyakori változtatását, valamint a hatékony adatkezelést és a modellekhez a megoldók hozzárendelését.

A GAMS modell kezelése a modellek szimbolikus, algebrai leírásán alapul, mert a szerzők véleménye szerint ilyen módon mind a felhasználók, mind pedig a gép számára pontos és könnyen érthető információkat szolgáltat. A nagyméretű modellek kezelését rugalmas indexelési lehetőség biztosítja.

A modellezési nyelv felépítése egyszerű, csak a minden gépen meglevő karakter készletet használja (a kis- és a nagybetűk egyaránt megengedettek). Az algebrai leírásban a szokásos algebrai és logikai műveleteken kívül egy új művelet is be lett vezetve. Ez a \$ művelet, amelynek definíciója szerint az $A \$ B$ úgynevezett feltételes kifejezésben az A kifejezést akkor és csak akkor tekintjük vagy számoljuk ki (a kifejezés tartalmától függően), ha a B kifejezés értelmezhető, vagy logikai kifejezés esetén az értéke igaz.

A GAMS modellezési nyelvnek relációs adatbáziskezelő rendszere van, amely támogatja egyrészt az adatok minél rugalmasabb kezelését (pl. akárhány dimenziós tömbök, akármilyen hosszúságú adatnevek szerepeltetése), másrészt a modell és az adatbázis leírásánál a kommentezési lehetőség kiterjesztésével a rendszer által generálható kutatási jelentés szerkesztését, valamint a modellek és a megoldók összekapcsolását.

9.3 A képernyőkezelés

A felhasználó a különböző szoftver rendszerekkel a képernyőkezelő rendszereken keresztül kerül kapcsolatba, melyek segítséget adnak a tudásábrázolásában és az ismeretek megszerzésében is. Ezért különleges figyelmet kell erre fordítani. Ha csak valamilyen nyomós ok nincs, akkor a képernyőkezelésnél építeni kell a legmodernebb szoftver technikákra, megkönnyítve ezzel a rendszer megtanulását és használatát.

Mint az előzőekből is kiderült, az ilyen bonyolult rendszerek felépítése, vezérlése és kezelése annyiféle különböző funkciót feltételez, ami teljes alapossággal egy IBM PC gépen nem megvalósítható. Ezért törekedni kell minden rendelkezésre álló lehetőség kihasználására.

Ebben a részben arról lesz szó, hogyan lehet a meglevő szoftverekre építve vizuálisan interaktív matematikai programozási rendszereket kialakítani. A szakértői rendszerek szoftverjeit vizsgálva szembeötlő, hogy míg eddig a különböző típusú

szoftver termékek egészen különböző formában jelentek meg, addig manapság erőteljes standardizálódás figyelhető meg. Ennek részben az az oka, hogy a felhasználók felé kialakult a szolgáltatások egy nemzetközi standardja, ami természetesen a hardver és a szoftver eszközök nagyarányú fejlődésével folyamatosan fejlődik. Nem lehet ezért hosszú évekre előre tervezni egy rendszert.

A jelenlegi lehetőségeket figyelembe véve a képernyő kezelésnél a WINDOWS rendszert lehet alapul venni. Ez a szokásos ablakkezelési technikán túlmenően alapja lehet vizuálisan interaktív matematikai programozási rendszereknek.

A matematikai programozási rendszerek alapvető jellemzője, hogy a rendszer legfontosabb funkciója, a feladat megoldása nem a felhasználó aktív közreműködésével történik, hanem az optimalizáló rutinok „fekete dobozként” működnek. Ez nem okozna gondot, ha garantálni lehetne a feladatok gyakorlati szempontból kielégítő megoldását. Sajnos ez még közepes méretű feladatokra sem igaz, ezért különösen a matematikai programozási rendszerek közvetlen használatánál van jelentősége a vizuális interaktivitásnak.

Ezért indultak meg az utóbbi években a kutatások az algoritmus animáció területén. Az algoritmus animációnál a fő cél vizualizálni a fejlesztő vagy a felhasználó számára fontos információkat. Így olyan összefüggésekre is fény derülhet, amire eredetileg senki sem gondolt.

Az algoritmus animáció minőségi változást jelent a rendszer intelligenciájában, hiszen a fejlesztő és a felhasználó számára is lehetővé válik az algoritmus jobb megértése, az algoritmus működésének a nyomon követése és szükség esetén interaktív módon a beavatkozás.

A fejlesztő kitapasztalhatja az algoritmus hangolás módozatait (pl. a paraméterek beállítása), észreveheti a kritikus pontokat (pl. mi lassú, mi numerikusan pontatlan, stb.) és megpróbálhat algoritmikus úton segíteni.

A matematikai programozáson belül, az algoritmus animációval kapcsolatos eredmények alsósorban a lineáris programozás területén vannak, ahol a feltételei mátrix vizuális ábrázolása módot nyújt a struktúra felismerésére, az oszlop kiválasztás vizuális segítésére, a vízszintes és függőleges blokk struktúra valamint, hatékony degeneráció elleni stratégia kialakítására.

Az algoritmus animáción belül megkülönböztetünk on-line és off-line animációt. Az on-line animáció két szerepe az események egyidejű figyelése és a szükséges esetekben a beavatkozás. Ez inkább a matematikai programozási rendszerek közvetlen felhasználását teszi hatékonná, míg az off-line animáció inkább az osztott szakértői rendszereknél hasznos. Az off-line animáció abban különbözik az on-line animációtól, hogy ez utóbbi esetben az algoritmus futásával kapcsolatos vizuális információkat tároljuk, majd később, esetleg egy másik gépen visszajátsszuk.

Az utóbbi 5 évben születtek meg az első szoftver eszközök az animációhoz. Az ezekkel kapcsolatos alapkövetelmények, hogy egyszerűen implementálhatóak legyenek, ne legyenek lassúak és ne függjenek az algoritmustól (lehetőleg a géptől se). A legismertebb talán a Balsa 2, amely a fentebb említett alapkövetelmények közül csak a géptől való függetlenséget sérti meg. A Balsa 2 alapfilozófiája, hogy algo-

ritmus animációnál egyetlen nézet kevés, több párhuzamos vetületet kell tekinteni, több ablak egyidejű megnyitásával. A különböző ablakokban ugyanannak a képnek lehet egy nagy totálja vagy egy kis részének a kinagyítása, ugyanannak a mátrixnak a gráfos alakja és az incidencia mátrixa. Ezenkívül természetesen szerepelhetnek hagyományos információk is a többi ablakokban (pl. különböző normák alakulása, az aritmitikai műveletek végzése során hogyan alakul a pontos jegyek száma, iterációs riport, stb.). Az algoritmus animációnál meg lehet tervezni az érdekes események ábrázolását és a megfelelő pontokon a programok elágazását.

Az eddigiekből is kiderült, hogy az algoritmusok intelligenciájának a növelésében nagy szerepe van annak, ha több ablakot lehet egyszerre nyitni és ezekben lényegileg összetartozó dolgok láthatók vagy párhuzamosan, különböző programok futnak (pl. ugyanazt a feladatot két különböző algoritmus oldja meg). Ez mutatja, hogy ha valamilyen kizáró ok nincs, akkor érdemes a WINDOWS vagy más hasonló tulajdonságokkal rendelkező rendszerrel összekapcsolni a matematikai programozási rendszert.

9.4 A matematikai programozási rendszer és az alkalmazói rendszer kapcsolata

A matematikai programozási rendszer modellezési nyelve, a könnyen megtanulható és kényelmesen használható GAMS kapcsolatot létesít a matematikai programozásban jártas szakember és a gép között. Azonban a végfelhasználó általában járatlan a matematikai programozásban, ezért biztosítani kell a megfelelő kapcsolatot a GAMS és egy valamilyen alkalmazói rendszer között.

Az ideális nyelv olyan lenne, amelyik bármely gyakorlati területen alkalmazható. Ezzel kapcsolatban történt előrehaladás, ugyanis 1979-ben megjelent a C nyelven megírt RITA, majd két évre rá a logikailag továbbfejlesztett változata, a ROSIE (The ROSIE Language Reference Manuel, 1981).

A ROSIE nyelv fő jellemzője az angolhoz nagyon közelálló szintaxisa és gazdag nyelvtana (pl. a nyelv a múlt időt ismeri). Ebből következik, hogy a nyelv nagyon könnyen olvasható, de nehezen megtanulható. A nyelv gazdagsága miatt lassú, és csak nagygépen futtatható. Ennek ellenére a nyelv létrejötté és alkalmazásának tanulságai nagy jelentőségűek lehetnek a továbbfejlődés szempontjából.

A ROSIE nyelv kialakulása óta három implementációja készült el, és egy folyamatban van. Ezek a következők:

- INTERLISP nyelven DECSYSTEM-20 gépen,
- VAX-INTERLISP nyelven VAX 11/780 gépen,
- INTERLISP-D nyelven XEROX SIP 1100 gépen és
- PORTABLE STANDARD LISP nyelven folyamatban van.

Mivel IBM PC gépekre jelenleg nem ismeretes olyan nyelv, amelyik minden döntési probléma esetén biztosítaná a mai elvárások mellett a modellezést és az adatbevitelt a terület-specifikus ismeretek alapján, ezért általában csak annyi mondható, hogy a szakértői rendszer pontos céljának ismeretében lehet részleteiben foglalkozni a kérdéssel.

Ha a döntési problémákat leszűkítjük az angol nyelvű irodalomban PDI (Production, Distribution, Inventory) problémáknak nevezett esetekre, akkor lehetőség

van a matematikai programozásban járatlan felhasználók számára is a szakterület-specifikus ismeretek alapján történő modellezésre és adatbevitelre. A PDI problémákról részletesen lehet olvasni a következő cikkekben és könyvekben:

- Bradley, Hax és Magnatti (1977),
- Glover, Jones, Karney, Klingman és Mote (1979),
- Glover és Klingman (1981),
- Helferich (1983),
- Kendrick, Meeraus és Alatorre (1983) és
- Klingman, Phillips, Steiger, Wirth, Padman és Krishnan (1987).

A PDI problémák esetén röviden arról van szó, hogy a döntési feladatok valamilyen termék előállításával, raktározásával és elosztásával kapcsolatosak. Ennek a leszűkítésnek azért van jelentősége, mert ezek a modellek ilyen általánosságban leírhatók kvalitatív módon, az elsődrendű logika segítségével. Így az általános modell leírásánál az egyes fogalmak közötti kapcsolatokat is figyelembe lehet venni.

Ennek köszönhetően egy konkrét döntési probléma esetén csak az általános és konkrét fogalmakat kell összekapcsolni, mert ez már indukálja az általános modell alapján a konkrét döntési probléma modelljét is. Ez a modell még természetesen nincs matematikai formában, azonban lehetőség van az automatikus átalakításra.

A lineáris programozásra vezető PDI problémákra készült, a fent vázolt elképzelések alapján a PDM és PM* modellezési nyelv, melyek a modellekről egy-egy GAMS file-t generálnak. Ezekről részletesen lehet olvasni Krishnan (1988/1) és (1988/2) cikkeiben.

Úgy gondoljuk, hogy az előbb említett lehetőségek inkább a jövő fejlesztési irányait mutatják. Jelenleg reális alternatívának a spreadsheet alapú rendszerek tűnnek.

10. Matematikai programozási taxonómia egy megvalósítása (Rövid felhasználói útmutatóval)

10.1 A matematikai programozási rendszer előzményei

Az intézetünknel elkészült matematikai programozási rendszer a következő munkák eredményein alapszik: A project egy korábbi fázisában a „Döntési eljárások egy módszertani taxonómiai rendszere” című problémafeltáró tanulmányban rendszereztek a döntéstámogatással és optimalizálással kapcsolatos matematikai diszciplínákat, egyúttal áttekintették a szakirodalomban található modellezési rendszereket az alkalmazott tudásábrázolási módszerek szempontjából. A levonható tanulságok felhasználásával a szakértői rendszerek tudásábrázolási technikáira építve kidolgoztuk a taxonómiai tudás számítógépes reprezentációs rendszerének alapjait. A taxonómiai rendszer számítógépes megvalósítása volt a következő lépésként kidolgozott CLASS-TAXOR szakértői rendszer. A módszertani taxonómiai rendszer a döntéstámogatási módszerek, modellek tudásbázisának szerepét játssza. A rendszer

tudáskeret (frame) alapú, az osztályozási alapfunkció és a bővíthetőség szempontjainak messzemenő figyelembevételével készült. Végül kiépítettük a rendszer megoldóbázisát, azaz a taxonómiai rendszerbe integráltunk számos konkrét matematikai programozási feladatot megoldó programrendszert.

10.2 A CLASS-TAXOR rendszer felépítése

A programrendszer az alábbi fő komponensekből áll:

1. A CLASS szakértői keretrendszerből, amely önállóan is használható, más rendszerek fejlesztésére is alkalmas (szűkített) frame alapú shell. A CLASS PROLOG nyelven íródott, tizenhét forrásmodulból áll, amelyek együtt kb. hatvan oldalnyi forrásprogramot jelentenek. A végrehajtható program mérete (tudásbázis nélkül) kb. 170 K, így egy kisebb alkalmazás futtatásához legalább negyed megabyte szabad tárhely szükséges. A shell számára a tudásbázist saját nyelvén az ún. frame nyelven kell megadni. E nyelv elemzőjét a rendszer tartalmazza. Hangsúlyozzuk, hogy a rendszer következtető gépe (interface engine) — bár PROLOG-ban íródott — alapvetően nem a PROLOG beépített visszalépéses technikáját használja. Ebből a szempontból tekintve a PROLOG-ot nem mint mesterséges intelligencia nyelvet, hanem mint általános célú programnyelvet használtuk.

2. A második komponens a taxonómiai rendszert leíró tudásbázis, ami hűsz frame nyelvű forrásfile alakjában található. A forrásmodulok szervezése olyan, hogy a főprogram (TAXOR.FRM) megfelelő pontjain inkludálódik a megfelelő szűkebb szakterületet reprezentáló tudásanyag. Emellett az egyes szakterületek (pl. statisztika, matematikai programozás, előrejelzés stb.) tudásanyaga önállóan is lefordítható és használható.

3. A harmadik komponens a rendszerbe integrált külső programok összessége, melyek a megoldóbázist alkotják. Jelenleg a MILP, az Expert Choice, a Condor-GDSS, a MULTPC és a GAMS megoldói a MINOS és a ZOOM érhetők el a rendszerből.

A következő pontban részletesen ismertetjük a CLASS rendszert, először a tudásábrázolás nyelvét, majd a program funkcióit.

10.3 A taxonómiai tudásábrázolás keretrendszere

A CLASS keretrendszer taxonómiákat, osztályozási feladatokat megvalósító szakértői rendszerek létrehozására és futtatására szolgál. Ebből adódóan a rendszert két szempontból tekinthetjük: a tudásbázist létrehozó fejlesztő, illetve az elkészült tudásbázissal konzultációt folytató felhasználó szemszögéből.

A felhasználó számára a CLASS-szal készült osztályozási rendszerek működését egy növényhatározó használatához hasonlíthatjuk. A program a besorolandó objektumot egy osztályozási fában helyezi el. Ennek során az osztályozási fa gyökerétől indulva a levelek felé lépkedünk, miközben a rendszer a vizsgált objektum tulajdonságaira vonatkozó kérdéseket tesz fel. A felhasználó a kérdésekre menükből való választással felel. A fa csomópontjaihoz rendelt osztályozási szabályok és az összes korábban kapott válasz kiértékelésével dönt a rendszer arról, hogy az osztályozási fában merre lehet továbblépni, azaz, hogy a vizsgált objektum mely szűkebb

osztályba sorolható be. Az egyes csomópontokban lehetséges külső eljárások aktivizálása. A rendszer lehetővé teszi az osztályozási fa megtekintését. Mivel a fa néhány száz, ezer csúcsot is tartalmazhat, ezért természetesen csak a fa egy kis részlete, egy választott csúcsból gyökerező 3-4 szint mélységű részfa jelenhet meg egyszerre a képernyőn. A fa vizsgálatokor vagy konzultáció közben a rendszer megmutatja a csomópontokhoz rendelt osztályozási szabályokat és a szabályokban szereplő ismert értékű változókat. A konzultáció során ismertté váló változók (a vizsgált objektum tulajdonságaira vonatkozó állítások) egy stack-re kerülnek, amit a felhasználó megtekinthet és módosíthat. Ily módon a konzultáció lényegében az osztályozási fa bejárását jelenti, amit a felhasználó a válaszaival interaktívan vezérel. A külső programok futtatása mellett a konzultáció eredménye mindazon csúcsok összessége, amelyek a fa bejárásakor a kiindulási pontból (általában a gyöktől) elérhetők. E pontokat a rendszer a kezdőpontból hozzájuk vezető úttal adja meg.

A tudásbázist létrehozó programozó számára a CLASS a népszerű Turbo nyelvekhez, Turbo Pascal-hoz és Turbo C-hez hasonló integrált, menüvezérelt fejlesztő környezetet biztosít. A tudásbázis kódolása a rendszer saját nyelvén, az ún. frame nyelven történik. Ezt a nyelvet — amely a frame rendszerek egy szűk speciális esetét valósítja meg — kifejezetten taxonómiák, osztályozási fák leírására terveztük. A frame nyelvű forrásprogramok előállítására legcélszerűbb a CLASS beépített editorában dolgozni, de bármely más szövegszerkesztő is használható. A CLASS esetében a fordítás a forrásprogramként leírt tudásbázisból egy memóriában tárolt (relációs) adatbázis kialakítását jelenti. A CLASS következtető gépe a konzultáció során ezen az adatbázison végez műveleteket. A fordításkor észlelt hibák jelzésének és kezelésének módja, hasonlóan a Turbo Pascal-hoz, segíti a hibák gyors, kényelmes javítását. A tudásbázisok fordítása additív művelet, azaz lehetséges, hogy egy már lefordított tudásbázis használata közben hozzávegyünk (fordítsunk) a rendszerhez valamely részterülettel behatóbban foglalkozó tudásanyagot. A tudásbázis lefordított állapotában lemezre menthető, és újból betölthető, így egy kész tudásbázis használatakor megtakaríthatjuk az egyébként nem túlságosan hosszú fordítási időt. A mentésnél a konzultáció során ismertté vált tények is megőrződnek, ezért ez a művelet alkalmas arra, hogy a konzultációt bármikor felfüggesztjük, és később a megszakítás előtti állapotából folytassuk.

A CLASS rendszer oly módon képes elindítani külső programokat, hogy azok futása közben ő maga nincs a memóriában. Ehhez a CLASS-t a `-class.bat` batch fájljon keresztül aktivizáljuk.

10.4 Taxonómiák megadása frame nyelven

A CLASS-ban készült szakértői rendszerek tudásbázisát a rendszer saját nyelvén, az ún. frame nyelven lehet leírni. A CLASS rendszer tartalmaz egy fordítóprogramot, amely a forráskódból interpretálható belső adatstruktúrát állít elő. Az alábbiakban a frame nyelv részletes ismertetésével foglalkozunk.

10.4.1 Osztályozási fa

A CLASS rendszer nyelvének elemi építőkevei az ún. frame-ek. A CLASS

osztályozási fák felépítésére alkalmas, a frame-ek az osztályozási fa pontjaival (node) azonosíthatók. Egy frame általános alakja a következő:

```
a_frame_neve(a_frame_megelőzőjének_neve)
<deklarációk>
{
<logikai kifejezés>
}
```

A fenti séma magyarázata előtt a következő általános megjegyzéseket tesszük: A frame nyelvben a szóköz az elválasztó funkciót tölt be, máskülönben megengedett a szabad írásmód. A programszövegben szereplő soremeléseknek, a program egy soron belüli elhelyezkedésének semmi jelentősége nincsen. A programban előforduló nevekről azt kell tudni, hogy egy név betűvel kezdődő, betűkből és számokból álló sorozat, ahol az aláhúzás karakter betűnek számít, a kis- és nagybetűk különbözők. A név hosszára nincs megkötés, a fordító a név teljes hosszát figyelembe veszi.

Az `a_frame_neve` karaktersorozat ezek szerint egy szabályos név, amivel a megfelelő frame definícióját bevezetjük. A `frame_ek` definíciójához nem használunk semmilyen kulcsszót, a fordító a szöveggörnyezet ismeretében feltételezi, hogy egy (kulcsszóként nem használatos) új név előfordulása új frame definíciójának kezdetét jelenti. A frame nevének egyedinek kell lenni, átdefinálás nem megengedett. Az újonnan létrehozott frame neve után zárójelek között szerepel a frame fabeli megelőzőjének neve. Ennek egy létező frame-nek kell lennie, előrehivatkozásra nincs mód. A rendszer automatikusan generálja a root nevű frame-et, amit a programok a fa gyökereként használhatnak. Frame-ek definíciójának sorozatával tehát megadhatunk egy fát, és mivel a fordító hibaüzenettel közbeavatkozna, fától különböző struktúra előállítása nem is lehetséges.

A frame deklarációs része után — amivel később foglalkozunk — kapcsos zárójelben egy logikai kifejezés áll. A frame-ekhez rendelt logikai kifejezések (osztályozási szabályok) vezérlik a fa bejárását. Egy vizsgált objektum osztályba sorolása a frame program végrehajtásával történik, ami viszont nem más, mint az osztályozási fa bejárása. A fa bejárása a következő szabály szerint történik:

Tegyük fel, hogy a fát bejárva eljutottunk az F nevű csomópontig (frame-ig).

1. Ekkor a rendszer kiértékeli F azon legbaloldalibb (a forrásprogramban legelőször definiált) fiánál lévő logikai kifejezést, amit eddig még nem vizsgált.
2. Ha az itt talált logikai kifejezés igaz értékű, akkor a rendszer bejárja az ebből a csúcsból gyökerező részét.
3. Ha F -nek maradt még nem vizsgált fia, akkor az eljárás az 1. pontnál folytatódik, egyébként visszalépünk F megelőzőjéhez.
4. Az eljárás a kiinduló pontba való visszaérkezéssel ér véget.

10.4.2 Logikai kifejezések

A frame program vezérlésében főszerepet játszó logikai kifejezések (osztályozási szabályok) atomi tényezők felett értelmezett Boole-formulák. A negáció operátorának jele a \sim karakter, amely megelőzi a hozzátartozó operandust. A konjunkció

operátorának $\&$, a diszjunkciónak a $|$ karakter a jele. Ez utóbbi két operátort a szokásos infix írásmóddal használjuk. Az operátorok precedenciája a szokásos: a negáció a legmagasabb, a logikai összedés (diszjunkció) a legalacsonyabb. A formulákban tetszőleges mélységig használhatunk zárójeleket. A Boole-formulák atomi tényezőinek helyén az alábbi objektumok állhatnak:

1. Logikai konstans:

Az igaz érték jele a t , a hamis értékű konstans jele az f karakter.

2. Property:

Ez a típus egy logikai változót jelent, amelynek igaz vagy hamis értéke a besorolandó objektumra vonatkozó valamely állítás érvényességét vagy érvénytelenségét fejezi ki. Definíálásának módja:

`property a_property_neve`

(A vizsgált objektumra vonatkozó állítás)

Egy property definíálását a property kulcsszó vezeti be. Ezután a property típusú változó neve áll, amelyre ugyanazok a formai követelmények érvényesek, mint a frame-ek nevére. A frame programban a property-re ezzel a névvel hivatkozhatunk, ez megfelel más programnyelvek globális változónév használatának. A property nevét követi a property által képviselt állítás szöveges megfogalmazása zárójelpár közé zárva. Ha a konzultáció során valamely logikai kifejezés kiértékelése közben szükség van egy property konkrét értékére, és ez az érték nem következtethető ki egyéb szabályból, akkor a rendszer a felhasználótól kér információt az állítás érvényességét illetően. A felhasználóval folytatott kommunikáció során az állítás mindig a szöveges formájában jelenik meg pl. a menükben. A property-k külső definícióként szerepelhetnek, nevük kötelezően egyedi, és definíálásuknak meg kell előznie a használatukat. Az előrehivatkozás programhiba.

3. Choice:

Ez a típus a property típusának egyszerű általánosítása. Olyan property-k együttesét jelenti, amelyek közül egyszerre legfeljebb egy igaz. Definíálásának formája:

`choice a_choice_neve`

{

`első_alternatíva(Első állítás)`

`második_alternatíva(Második állítás)`

`...`

`n.edik_alternatíva(N-edik állítás)`

}

A definíciót a choice kulcsszó vezeti be, ezt követi a choice neve, majd kapcsos zárójelpár között a choice alternatíváinak felsorolása. Az alternatívák ugyanúgy a vizsgált objektumra vonatkozó állításokat fejeznek ki, mint egy property, ezek az állítások azonban kizáróak. A frame programban egy choice alternatívára a C-ből kölcsönzött jelölésmóddal, a struktúra tagokhoz hasonlóan hivatkozhatunk. Az alábbi kifejezés

`a_choice_neve.n.edik_alternatíva`

például logikai változóként szerepelhet egy Boole-formulában, és értéke az „ N -edik állítás” érvényességét fejezi ki. A konzultáció közben a rendszer a choice értékét is megkérdezheti a felhasználótól, ekkor az alternatívákhoz tartozó állítások szövege együtt jelenik meg egy menüben. A property típushoz hasonlóan a choice is külső definícióként szerepelhet, a choice nevének egyedinek kell lennie (de nem az alternatíva neveknek). Előrehivatkozásra itt sincs mód, a choice definiálásának meg kell előznie a használatát.

4. Frame:

Atomi tényezőként szerepeltethetjük a Boole-formulákban egy korábban definiált frame nevét. Ennek pontosan ugyanaz a hatása, mintha az illető frame-hez rendelt Boole-formulát zárójelekkel körülfogva beírnánk a frame nevének helyére. Ez egyfajta szubrutinhívásnak felel meg a frame nyelvben.

5. Beépített eljárások:

`open('string')`: Megnyitja és írásra kijelöli a string nevű file-t. Értéke igaz, ha a megnyitás sikeres, hamis, ha nem.

`write('string')`: Beírja string-et az írásra megnyitott file-ba. Értéke mindig igaz.

`close`: Lezárja az előzőleg megnyitott file-t. Mindig igaz.

`edit('string')`: Felajánlja editálásra a string nevű file-t. Ha a file létezik, akkor edit értéke igaz függetlenül attól, hogy akarunk-e editálni vagy nem. Edit értéke hamis, ha a file nem létezik.

`batch('string')`: Felajánlja elindításra a string nevű batch file-t. Értéke akkor és csak akkor igaz, ha a file létezik.

10.4.3 Szabályok

Gyakran előfordul, hogy a vizsgált objektum tulajdonságait kifejező állítások között valamilyen belső összefüggés létezik. Bizonyos property-k és choice alternatívák értékének ismeretéből már következtetni lehet más állítások érvényességére. Ilyenkor kíváncsi, hogy ezeket a következtetéseket a rendszer önállóan tegye meg, hiszen felesleges kérdezősködéssel csak terhelné a felhasználót, aki általában nincs is tudatában az említett belső összefüggéseknek, és így könnyen ellentmondásos válaszokat adhatna. E probléma megoldására a CLASS rendszer tartalmaz egy Horn-clause-os szabály-kiértékelő mechanizmust. A szabályokat a következő alakban lehet megadni:

$(A \rightarrow X)$, vagy

$(A \rightarrow \sim X)$, vagy

$(A \& B \& C \rightarrow)$ stb.

A szabályok a külső definíciók helyén állhatnak, a szabályokat leíró clause-t zárójelek közé tesszük. A szabályban szereplő A, B, C, X szimbólumok korábban definiált property-k vagy choice alternatívák lehetnek. A következtést kifejező nyíl a „mínusz” és a „nagyobb” karakterekből tevődik össze. A szabály bal oldala nem lehet üres. A bal oldali szorzatban szereplő tényezők száma nics korlátozva, de nem használható

a \sim és $|$ operátor. A szabály jobb oldala lehet üres, vagy állhat egy (esetleg negált) property-ből vagy choice alternatívából. Felhívjuk a figyelmet arra, hogy a példák közt szereplő első két szabályból együttesen már A hamis értéke következik, pedig A nem szerepel egy szabály jobb oldalán sem.

10.4.4 A logikai kifejezések kiértékelése

Miután minden építőkövet ismerünk, áttekinthetjük az osztályozási fa pontjaihoz rendelt Boole-formulák kiértékelésének pontos rendjét. A Boole-formulák kiértékelése balról jobbra haladva történik. Eközben, ha egy többtényezős logikai szorzat valamely tényezője hamisnak bizonyul, akkor az egész szorzat azonnal hamis értéket kap, és a szorzat még hátralévő tényezőit a rendszer már nem vizsgálja. Hasonlóan, ha egy többtagú logikai összeg egy tagja igaz, akkor az egész összeg is igaz, és a (jobbra) még hátralévő tagokat a rendszer nem vizsgálja. A formulák kiértékelésének ezt a mechanizmusát a frame programok vezérlésátadási szerkezetek létrehozására használják.

A choice-ok, property-k kiértékelése a következő szabály szerint zajlik. Ha a változó értéke ismert, például azért, mert a felhasználó már megadta, akkor semmi teendő nincs, a program számol az ismert értékkel. Ha a változó még nem ismert, akkor a rendszer megvizsgálja, hogy az előző pontban tárgyalt szabályokból kikövetkeztethető-e a változó értéke. Ha igen, akkor a kapott értéket bejegyzí az ismert tények közé, és ugyanúgy jár el, mint az előző esetben. Ha a szabályok vizsgálata sem vezetett eredményre, akkor a rendszer a felhasználótól kér információt. Ez úgy történik, hogy az éppen vizsgálat alatt lévő property vagy choice alternatívák állításainak szövege megjelenik a képernyőn egy menüben, amiből a felhasználó választ. Választás helyett a felhasználó megnyomhatja az Esc billentyűt, ennek hatására az éppen kérdezett változók nem kapnak értéket, az aktuálisan kiértékelés alatt álló osztályozási szabály vizsgálata megszakad, és az osztályozási fa bejárása úgy folytatódik, mintha a kiértékelés hamis eredményre vezetett volna.

10.4.5 Deklarációk

A frame logikai kifejezésében használt property-ket, choice-okat és frame-eket deklarálni kell. A deklaráció célja a fordítás biztonságosabbá tétele a gépelési hibából adódó véletlen névegyezés, típus tévesztés és más hasonló hibák elleni védelemmel. A deklarációk a frame fejrésze és a logikai kifejezés kezdő kapcsos zárójele között helyezkednek el. Alakjuk:

```
property első_property, ..., n_edik_property;
choice első_choice, ..., n_edik_choice;
frame első_frame, ..., n_edik_frame;
```

A deklarációs kulcsszó után a deklarálendő elemek vesszővel elválasztott sorozata áll, amit pontosvessző zár le. Akármelyik deklarációs kulcsszó szerepelhet többször is. A felesleges deklaráció nem hiba, de a deklarált objektumnak létezőnek, azaz korábban már definiálnak kell lennie. Choice-ok esetében csak a choice nevet kell deklarálni, az alternatívák nevét nem.

10.4.6 BNF szintaxis

Az előző pontokban ismertetett frame nyelv szintaxisát tömören az ún. BNF-fel (Backus–Naur Form) foglалhatjuk össze.

```

<program> ::= external_def [ <program> ].
<external_def> ::= <property>
| <choice>
| <rule>
| <frame>.
<property> ::= property <P_name> ( <statement> ).
<choice> ::= choice <C_name> <alt_list>.
<rule> ::= ( <left_side> -> <right_side> ).
<frame> ::= <F_name_1> (<F_name_2>) <dec_list> <condition>.
<P_name>, <C_name>, <F_name>, <A_name> ::=
    Betűvel kezdődő, betűkből és számokból álló sorozat, az aláhúzás is betű,
    a kis- és nagybetűk különbözők.
<statement> ::= Zárójelet nem tartalmazó string.
<alt_list> ::= <A_name> (<statement>) [<alt_list>].
<R_boolean> ::= = <P_name>
| <C_name>.<A_name>.
<C_boolean> ::= (<R_boolean>
| <F_name>.
<left_side> ::= = <R_boolean> [ & <left_side> ].
<right_side> ::= [ [~] <R_boolean> ].
<dec_list> ::= [ <declaration> [<dec_list>] ].
<declaration> ::= property <P_name_list>;
| choice <C_name_list>;
| frame <F_name_list>;.
<name_list> ::= [ <name> [, <name_list>] ].
<condition> ::= <term> [ | <condition> ].
<term> ::= <factor> [ & <factor> ].
<factor> ::= t
| f
| <C_boolean>
| open(''string'')
| write(''string'')
| close
| edit(''string'')
| batch(''string'')
| exit
| (<condition>)
| ~<factor>.
```

A nyelv további bővítésének egy lehetséges útja a BNF-ben <factor>-nak nevezett objektum további alternatíváinak létrehozása. Az atomi logikai tényezők kiértékelésének mellékhatásaként valósítható meg egy sor művelet. Például aritmetikai kifejezések közti reláció — mint logikai atom — kiértékelése közben elvégezhetők bizonyos számítások. Logikai érték rendelhető bizonyos relációk, pl. I/O műveletek, démonok végrehajtásának eredményéhez. További közeledést jelent a teljes frame rendszerekhez a változók érvényességi tartományának struktúrálása. Ha a változók definiálása a frame-ekhez köthető (jelenleg minden változó root-ban definiált, azaz globális), akkor egy változót a definiálási pontjából elérhető pontokban tekinthetünk létezőnek.

10.5 A taxonómiai rendszer funkciói

Ha a class rendszerből külső programokat is aktivizálni akarunk, akkor indítsuk a — class.bat file-on keresztül. Paraméterként megadható annak a lefordított tudásbázisnak a neve, melyet a rendszer automatikusan betölt (kiterjesztéssel együtt kell a nevet megadni, pl. taxor.kba). Indítása után a CLASS rendszer egy pull-down menüvel jelentkezik be, amelyből a rendszer fő funkciócsoportjai érhetők el. Az alábbiakban a rendszer funkcióinak használatát ismertetjük a manürendszernek megfelelő csoportosításban.

10.5.1 Konzultáció

A funkciócsoport a pulldown menü (főmenü) Consult ágának kiválasztásával érhető el. A kiválasztás kétféleképpen történhet. A kurzormozgató billentyűkkel, majd az Enter billentyű lenyomásával, vagy a menücím intenzív módban megjelenített karakterének beütésével. A Consult menüben a következő funkciók választhatók ki: Classify, Taxonomy, View facts, Remove facts.

1. Classify:

Ebben a menüágban a lefordított frame program futtatása, az osztályozás történik. Aktivizálódik a Consultation és a Classification feliratú ablak. Előbbiben megjelenik egy kérdés, mely a felhasználótól az osztályozási fa azon csúcának nevét kéri, ahonnan a felhasználó indítani szeretné az osztályozást. A megfelelő csomópont (frame) nevének begépelése, majd az Enter lenyomása után kezdődik az osztályozás.

Egyéb lehetőségek: Ha az osztályozást a gyökérből akarjuk kezdeni, nem kell begépelni a root nevet, elegendő csak az Enter lenyomása. A kérdőjel billentyű lenyomására a rendszer a Consultation ablakban felsorolja a létező csomópontok nevét, majd megismétli a kérdést. Esc lenyomásával az osztályozás elindítása helyett kiléphetünk a menüágból.

A konzultáció során a rendszer a Classification ablakban mindig mutatja, hogy éppen melyik frame logikai kifejezésének kiértékelése folyik, ugyanitt megjelenik a kiértékelés eredménye is. Az osztályozandó objektum még nem ismert tulajdonságaira vonatkozóan a rendszer a felhasználótól kér információt. Az információ menüből való választással történik: a kurzort a helyes állításra kell vinni, és lenyomni az Enter billentyűt.

Egyéb lehetőségek: A tízes funkcióbillentyű (F10) lenyomásával aktivizálható az Explanation ablak, amelyben megtekinthető az aktuálisan vizsgált Boole-formula szöveges alakja, és a formulában szereplő ismert tényezők értéke. Az Explanation ablakból az Esc vagy F10 billentyűvel léphetünk ki, ami után a konzultáció folytatódik. A rendszer kérdésre a menüből való választás helyett a felhasználó felelhet az Esc billentyű lenyomásával. Ekkor a kérdezett változók nem kapnak értéket, a vizsgált Boole-formula kiértékelése megszakad, és a frame program végrehajtása úgy folytatódik, mintha a kiértékelés eredménye hamis lett volna.

A programfutás eredménye mindazon frame-ek összessége, amelyek a kezdőpontból elérhetők. Ezeket a rendszer mindig a kezdőpontból hozzájuk vezető teljes út felsorolásával mutatja a Consultation ablakban. Egy-egy megoldás megtalálása után Esc leütésével befejezhetjük az osztályozást, vagy akármely más billentyűt lenyomva további megoldásokat kereshetünk.

2. Taxonomy:

E menüágban a rendszer megjeleníti az osztályozási fát. Mivel a fa akár ezer csúcsból is állhat, egyszerre csak a fa egy kis részlete, egy választott csúcsból gyökerező 3–4 szint mélységű részfa látható a képernyőn. A Taxonomy funkcióba való belépés után először egy menü jelenik meg, amelyben a fa összes csúcsa fel van sorolva. Ebből a menüből választható ki a képernyőn ábrázolt részfa gyökere. A kurzor-mozgató billentyűkkel tologathatjuk a fát a képernyőn, illetve kiválaszthatjuk az egyes csomópontokat. Egy csomópont kiválasztása után az Enter lenyomásával a fa további, a kiválasztott pontból gyökerező részfája jeleníthető meg. Az eljárás akárhányszor (a memória beteltéig) ismételhető, visszatérés az Esc billentyűvel a korábbi állapotokon sorban visszafelé haladva lehetséges. A kiválasztott csomópont logikai kifejezése is vizsgálható az Explanation ablakban — ugyanúgy, mint osztályozás közben — F10-zel.

3. View facts:

A konzultáció közben a rendszer kérdéseire adott válaszok, illetve a rendszer által kikövetkeztetett tények ismertté válásuk sorrendjében egy stack-re kerülnek. A View facts funkcióban ez a stack egy menüben vizsgálható. A kiválasztott menütétel és a stack e feletti része a Del billentyű lenyomásával eltávolítható. A funkció alkalmas arra, hogy a hibás válaszokat „visszavegyük”, majd a Classify funkcióba visszalépve a hibás válasz helyétől a hiba kijavításával folytassuk az osztályozást.

4. Remove facts:

E funkcióval az összes válasz és következtetés egyszerre törölhető, ezután az osztályozást teljesen előlről kezdetűk.

10.5.2 Műveletek a tudásbázissal

Ez a funkciócsoport a Knowledge menüág kiválasztásával aktivizálható. A csoport a frame programok fordítását, a lefordított tudásbázis törlését, betöltését és mentését tartalmazza.

1. Compile:

A menüág kiválasztása után a képernyőn megjelenik az aktuális alkönyvtár `.frm` kiterjesztésű állományainak listája. A frame nyelvű forrásprogramok szokásos kiterjesztése `.frm`, de más file-maszk is megadható F4 lenyomása után. A kívánt forrásprogram kiválasztását követően Enter leütésével elkezdődik a fordítás. Ha a fordítás közben hiba lép fel, akkor a rendszer automatikusan editor üzemmódba kerül, a kurzor a hibás file első értelmezhetetlen karakterére áll, az alsó sorban pedig a hiba jellegére utaló rövid üzenet olvasható. A hibát azonnal javíthatjuk, a kurzor első mozdítására a hibaüzenet eltűnik. (Az editálással kapcsolatban lásd a megfelelő funkció leírását.) A programfordítás additív művelet, ez azt jelenti, hogy különböző forrásfile-okat egymás után lefordítva a bennük lévő tudásanyag összeadódik. Lehetséges tehát, hogy konzultáció közben hozzávegyük a rendszer tudásbázisához valamely szakterület részletesebben kidolgozott anyagát.

2. Erase kbase:

Hatására a teljes tudásbázis törlődik, a rendszer alapállapotba kerül.

3. Load kbase:

Egy előzőleg elmentett tudásbázist tölt be lemezről. A képernyőn megjelenik az aktuális alkönyvtár `.kba` kiterjesztésű file-jainak listája. A megfelelő file kiválasztása után Enter hatására a file betöltődik. A memóriában a betöltés előtt lévő tudásanyag elvész, a művelet a rendszert pontosan abba az állapotba hozza, mint amiben a betöltött tudásbázis kimentésekor volt.

4. Save kbase:

A lefordított tudásbázist szabályokkal, ismert tényekkel stb. együtt a lemezre menti. Egy külön ablakban bekéri a létrehozandó lemezfile nevét, ha kiterjesztést nem adunk meg, akkor a `.kba` kiterjesztést tételezi fel.

10.5.3 Forrásprogram editálása

Az Edit menüág aktivizálásakor megjelenik a képernyőn az aktuális alkönyvtárban található `.frm` kiterjesztésű file-ok listája. A frame programok szokásos kiterjesztésére `.frm`, de más típusú file-okat is editálhatunk, ehhez nyomjuk le az F4-et, majd adjuk meg a kívánt új file-maszkot. Az editálni kívánt file kiválasztása után Enter-t leütve belépünk az editorba. A CLASS rendszer szövegszerkesztője egy normál Turbo editor, amelyben F1-gyel egy komplett help rendszer hívható elő, ezért az editor funkcióinak részletes ismertetésére nincs szükség. A szerkesztőből F10-zel vagy Esc-pel léphetünk ki. Az F10-es kilépésnél a módosított file-t a rendszer automatikusan a lemezre menti. Az Esc billentyűvel való kilépéskor a rendszer mielőtt veszni hagyná az esetleges módosításokat, megerősítést kér a felhasználótól.

10.5.4 A rendszer konfigurálása

A Setup funkció lehetővé teszi a felhasználó számára, hogy a CLASS rendszer ablakainak többségét átszínezzé, az ablakok helyét és méretét megváltoztassa. Színezéskor a megjelenő színmintából két színekombinációt kell kiválasztani, az első az ablak belsejének, a második az ablak szegélyének az új színe lesz. Ha a keret színére fekete alapon fekete kombinációt választunk, akkor az ablaknak nem lesz

kerete. Az ablak méretének és helyének változtatása a kurzormozgató és a shift billentyűk használatával történik.

10.5.5 Rendszerhívás, kilépés

A Quit menüág OS shell funkciója révén a felhasználó elindíthat egy új command.com parancsértelmezőt. Ez lehetővé teszi, hogy a rendszer használatát ideiglenesen megszakítva a maradék memóriában DOS programokat futtasson. A CLASS rendszerbe az exit DOS paranccsal lehet visszatérni.

A Quit menüág Exit funkciójával végleg elhagyjuk a CLASS rendszert. E funkció végrehajtása előtt a rendszer megerősítést kér a felhasználótól.

10.6 A TAXOR matematikai programozási rendszer

A TAXOR szakértői rendszer tudásbázisa húsz frame nyelvű file alakjában található. A forrásmodulok szervezése olyan, hogy a főprogram (TAXOR.FRM) megfelelő pontjain inkludálódnak az egyes részterületek tudásanyagát tartalmazó file-ok.

STATIST.FRM statisztika

SZIMUL.FRM szimuláció

QUEUE.FRM sorbanállás

ELORE.FRM előrejelzés

DANAL.FRM döntésanalízis

OPTIMAL.FRM inkludálja az alábbi file-okat:

HALOZAT.FRM hálózati optimalizálás

TABLA.FRM táblázatkitöltési modell

MATPROG.FRM matematikai programozás

DINA.FRM dinamikus modell: készletezés stb.

EGYENS.FRM egyensúlyi modell: játékelmélet stb.

A felsoroltakon kívül további néhány segédfile biztosítja, hogy az egyes szakterületek tudásanyaga önállóan is lefordítható, és így használható legyen. A teljes rendszer fordítása néhány percre tart, ezért érdemes a tudásbázist lefordított állapotban is megőrizni a főmenü Knowledge ágának Save kbase funkciójával. A program elindítása célszerűen a

`-class taxor.kba`

paranccsal történhet. Ez biztosítja, hogy a külső programok lefutása után a CLASS rendszer visszatöltődjön. Természetesen szükséges, hogy az elindítani kívánt külső programok rendelkezésre álljanak.

11. Összefoglalás

A hatékony döntéstámogatáshoz elengedhetetlen azoknak az eszközöknek egy-egy keretbe foglalása és ellenőrzött, felhasználóbarát módon való rendelkezésre

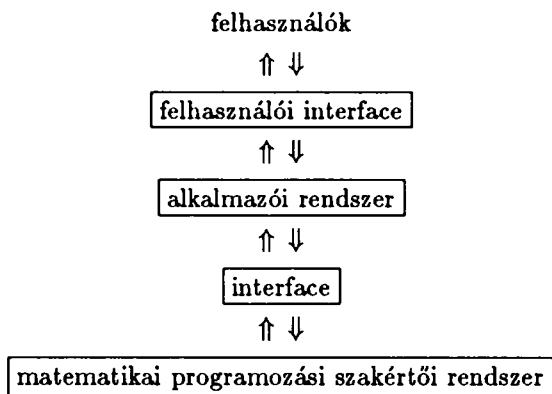
bocsátása, melyek a már jól specifikált problémák tényleges megoldására használhatók. A rendszernek tartalmaznia kell mindazokat az egységeket, amelyek szükségesek ahhoz, hogy a végfelhasználó a saját szakterületének terminológiájával fogalmazhassa meg a feladatot. Ehhez elsősorban a szóhajóvő feladatok modelljeihez megtervezett ismeretbázisra van szükség. Az ismeretbázisnak tartalmaznia kell az általános fogalmakat, fogalomtípusokat és a közöttük fennálló összefüggéseket. Ezeket az ismeretbázisban felismerhető és azonosítható formában kell reprezentálni. Az ismeretbázist továbbfejleszthetőnek képzeljük, ezért a kezdetben megadott modelleket tartalmazó ismeretbázis az alkalmazások során tovább épülhet, ha a felhasználó olyan feladatot fogalmaz meg, amelynek modellje nem szerepel az ismertbázisban.

Gyakorlati, üzemszerű felhasználás esetén e megközelítés jelentősége vitathatatlan. A döntéshozók, különösen a mai körülmények között, nem építhetnek merreven előre rögzített modellekre. Másrészt az sem elviselhető, ha egy új, a gyorsan változó körülményeknek megfelelő modell felépítése vagy adaptálása az új körülményekhez hosszabb időt vesz igénybe, mint amennyi a döntéshez rendelkezésre áll. A feladatazonosítás és a modellépítés ciklusának követnie kell a változásokat.

Ilyen típusú rendszereknél a felhasználónak nem kell feltétlenül ismernie a megoldandó feladatnak megfelelő matematikai programozási modellt és nem is kell azt megfogalmaznia, sőt azt sem kell tudnia, hogy a rendszer a válasz generálásához milyen modellt használ. A rendszer a felhasználóval folytatott interjú során maga választhatja ki a megfelelő modellt. Azonban megkötés az, hogy a modell csak matematikai programozási modell lehet. Természetesen más típusú modellek esetére hasonló rendszert lehet tervezni.

A szakértői rendszerek képesek a 2. fejezetben tárgyalt, döntéstámogatási modellalkotásra és feladatmegoldásra jellemző mind a nyolc lépést hatékonyan segíteni.

Az általános struktúra a 2. ábrán látható.



2. ábra

A 2. ábrán látható, hogy a legalsó szint a matematikai programozási szakértői rendszer, amelynek két fő felhasználási módja lehet:

- alkalmazói rendszeren keresztül és
- közvetlenül.

Az első a matematikai programozásban kevésbé járatos gyakorlati szakemberek számára kell, hogy biztosítsa a fellépő feladatok megoldását. A közvetlen felhasználás egyrészt a matematikai programozásban otthonosan mozgó, esetleg más területeken dolgozó specialisták számára, másrészt a rendszer fejlesztői és karbantartói számára kell, hogy biztosítsa a rendszerhez való kényelmes és közvetlen hozzáférést.

Bizonyos típusú döntési problémák megoldásánál hosszú idő óta alkalmazzák sikeresen a lineáris programozást. Azonban a legtöbb döntési probléma, amelyik megfogalmazható matematikai programozási feladatként, nem írható fel, mint lineáris programozási feladat. Ezért a szakértői rendszerek technikájának szélesebb körű elterjedéséig, a matematikai programozásban járatlan felhasználók számára nagyon kényelmetlenné tették az ilyen típusú feladatok megoldását a következő lépések:

- a megoldandó feladathoz generálni kellett egy speciális adatfile-t,
- a megoldandó feladathoz ki kellett választani a megfelelő megoldó algoritmust, majd az algoritmus alapján a megfelelő programot,
- a megoldó programokkal kapcsolatos dokumentációkat tanulmányozni kellett,
- mivel a megoldó programok általában FORTRAN nyelven íródtak, ezért a használatukhoz szükséges tömbdeklarációkat és a különböző paraméterátadásokkal kapcsolatos tennivalókat el kellett végezni,
- a különböző megoldó programokhoz különböző formában (esetleg különböző programnyelven) kellett megírni a probléma célfüggvényét és a feltételi függvényeket, valamint ezek gradienseit kiszámoló szubrutint,
- meg kellett határozni a gyakorlati problémának leginkább megfelelő eredmény file-t és
- a matematikai programozási feladat eredményeit át kellett fogalmazni a felhasználó fogalomköréhez közelálló formába.

A matematikai programozási szakértői rendszerek jelentősége abban áll, hogy használatuknál fölöslegessé válnak ezek a lépések, sőt a szakértői rendszerek minél pontosabban megfogalmazandó céljától függően még plusz szolgáltatások is beépíthetők (pl. analitikus és numerikus gradienst is számol, algoritmus animációt végez, stb.).

A matematikai programozási szakértői rendszerek működésére az a jellemző, hogy minden a rendszer által generált probléma bekerül a modellbázisba, ahonnan bármikor előhívható akár módosítás, akár egyéb célból vagy kitörölhető onnan. A matematikai programozási rendszerek hatékonyságát nagymértékben befolyásolja, hogy a szimbolikus formában megadott modellek kezelése és vizsgálata milyen fejlettségi szinten történik. A matematikai programozási szakértői rendszer a megadott probléma megoldására javasol egy megfelelő megoldó programot (lehetőség szerint teljesen automatikusan) és ha a felhasználó igényli, generál egy komplett forrásprogramot. Ennek az az előnye, hogy a generált programot a szakértői rendszer nélkül is lehet használni, a felhasználó által írt rutinokkal ki lehet egészíteni és így a rendszer egyben szoftver fejlesztő eszközként is működtethető. (Természetesen ezeket a változtatásokat a rendszer használata közben is meg lehet csinálni.)

Egy teljesen új modell beillesztéséhez elengedhetetlenül fontos, hogy az új modell adatstruktúrája illeszkedjen a rendszer adatstruktúrájához, ezért a matematikai programozási rendszereknek fejlett adatbázis kezelő rendszerrel kell rendelkeznie.

Ha bármely, a feladatban szereplő függvénynek és változónak tetszőlegesen hosszú neve lehet, valamint a rendszer a kísérő szövegek kezelésre is fel van készítve, akkor a megoldásról problémafüggő eredményt lehet nyerni, amely a probléma részletes megoldásán kívül, a megoldáshoz vezető út indoklását is tartalmazza.

Ha a rendszert képessé tesszük a feladatok megoldása során szerzett tapasztalatok összegyűjtésére és a tanulásra, akkor a rendszer ismeretbázisa és az intelligenciája folyamatosan fejleszthető. Így a rendszer működése a tapasztalatok növekedésével párhuzamosan egyre megbízhatóbbá és jobbá válik.

Ha a program a feladatot nem tudja megoldani, akkor egy szabályrendszer által vezérelt hibaanalízis adhat magyarázatot és tehet javaslatot a hiba okának a leküzdésére. Mivel a modellek osztályozása és a megoldó programok kiválasztása is egy szabályrendszer segítségével történik, ezért a rendszernek szüksége van lehetőleg minél jobban struktúrált, jól működő szabályrendszerre is.

A matematikai programozási rendszereknek kényelmesen kezelhető, könnyen megtanulható modellezési nyelvvel kell rendelkezniük, amelynek a segítségével modellezni lehet az aktuális problémát, kérni a már korábban megalkotott modelleket vagy a modellekhez tartozó adatok beolvasását, ezek bármelyikének a módosítását, törlését, a probléma editálását és megoldását, az egyes problémákhoz tartozó modellek és adatbázisok egymásba való transzformálását, a problémáról és a probléma megoldásáról készülő jelentés generálását, a megoldó program futásának megállítását, a rendszerről, a matematikai programozási problémákról vagy a megoldó algoritmusokról információt. Ezenkívül lehetővé kell tenni az operációs rendszer parancsainak a bevitelét is.

Ebből a rövid összefoglalásból is látszik, hogy egy matematikai programozási rendszer bonyolult, egyenként is nagyon nehezen megvalósítható és nehezen összekapcsolható egységekből áll. Ezért a komplett matematikai programozási rendszerekhez szükséges eszköztár rövid időn belüli hazai kifejlesztése az egészet tekintve irreális, bizonyos komoly hagyományokra épülő részeiben azonban nem.

Az MTA SZTAKI Operációkutatás és Döntési Osztályán kifejlesztettünk egy matematikai programozási modellkezelő rendszert, amelynek modellezési nyelve a GAMS (General Algebraic Modeling System) és amelyben a modellek taxonómiai osztályozása is megtörténik. A taxonómiai fa több, mint 500 csúcsot tartalmaz, a szabályok száma nagyobb, mint 500 és a megoldóbázis az ismertebb megoldók közül tartalmazza a ZOOM, az XMP, a MINOS és a MILP megoldókat. Az itt szereplő eredmények a továbblépés elméleti alapját jelentik.

A cikkben megkíséreltük egy olyan keretrendszer tervét összeállítani, amelyik a matematikai programozási szakértői rendszerek minden lényeges elemét tartalmazza. Így egy konkrét munka során pontosan lehet tudni, hogy milyen elemekből kell a rendszert felépíteni, amely természetesen a rendszer céljától függően más és más lehet.

Komplett matematikai programozási rendszerek kifejlesztését megelőzve, érdeemes már a jelenlegi munkák kereteiben olyan modulokat létrehozni, amelyek figyelembe veszik a piac igényeit, kompatibilisek lesznek a későbbiekben elkészülővel, s így különösebb változtatás nélkül összeépíthetők.

A főbb alkalmazási területek a következők lehetnek:

- speciális döntési problémák (vagy probléma osztályba tartozó feladatok) megoldására megfelelő számítógépes környezet biztosítása,
- speciális döntési problémák esetén különböző, a feladat természetéhez egyre jobban simuló modell variációk kipróbálása,
- különböző feladat osztályokon, a matematikai programozási algoritmusok tesztelése,
- numerikus tapasztalatok összegyűjtése a feladatok megoldásáról és
- a matematikai programozás oktatásához (modell építés, algoritmus kiválasztás, stb.) jól használható eszköz biztosítása.

Végezetül szeretnénk köszönetet mondani „A számítógépes döntéstámogatás új generációs eszközei” című project többi résztvevőjének, Csáki Péternek és Maros Istvánnak a munka során tanúsított jó együttműködésért és a hasznos tanácsokért.

IRODALOM

- [1] ANDRIOLE, S. J., *Microcomputer decision support systems: design, implementation and evaluation* (North-Holland, 1986).
- [2] ASZALÓS JÁNOS, „A tudásbeszerzés módszerei”, *Információ Elektronika* 1-2 (1990), 46–66.
- [3] BANERJEE, S. and BASU, A., „A knowledge based framework for selecting management science models”, *IS Technical Report TR-89-41* (University of Maryland at College Park, Maryland, 1989).
- [4] BAZARAA, M. S. and SHETTY, C. M., *Nonlinear programming, theory and algorithms* (John Wiley & Sons, 1979).
- [5] BERNAU, H., „Upper bound techniques for quadratic programming”, in *Survey of mathematical programming* (Prékopa A., ed.) (Akadémiai Kiadó, Budapest, 1979), 347–356.
- [6] BERNAU, H., „Ein Schnittebenenverfahren für das lineare Komplementaritätsproblem”, in *Methods of operations research* 45 (Verlagsgruppe Athenaum/Hain, 1983), 35–46.
- [7] BERNAU, H., „Ein neues sequentielles quadratisches Optimierungsverfahren”, *Seminarberichte Humboldt-Universität Berlin, Sektion Mathematik* 4 (1987), 28–44.
- [8] BINBASIOGLU, M. and JARKE, M., „Domain-specific DDS tools for knowledge-based model building”, *Working Paper*, CRIS # 97, GBA # 85–89 (New York University, 1985).
- [9] BÍRÓ, M., „Graphical representations in DSS and resource constrained scheduling”, *MTA SZTAKI Report* (18/1989).
- [10] BÍRÓ M. és CSÁKI P., „Döntési feladat-azonosítást támogató interjú-rendszer”, *Tanulmány* (OMFB-EKPI, Budapest, 1988).
- [11] BÍRÓ, M., MAYER, J., RAPCSÁK, T. and VERMES, M., „On building mathematical programming expert systems”, *Proceedings of the II. Conference on Artificial Intelligence* (Budapest, 1991), 155–163.
- [12] BISSCHOP, J. and MEERAUS, A., „On the development of a general algebraic modeling system in a strategic planning environment”, *Mathematical Programming Study* 20 (1982), 1–29.
- [13] BODILY, S. E., „Spreadsheet modeling as a stepping stone”, *Interfaces* 16 (1986), 34–52.
- [14] BOKOR J., TANYI M. és GÁSPÁR P., „Szabálybázisú szakértő rendszer fejlesztés INSIGHT 2+ környezetben”, *OTKA Tanulmány* (MTA SZTAKI).

- [15] BONCZEK, R. H., HOLSAPPLE, C. W. and WHINSTON, A. B., „Foundations of decision support systems”, *Academic Press* (1981).
- [16] BRACHMAN, R. J. and LEVESQUE, H. J., „A fundamental tradeoff in knowledge representation (Revised version)”, in *Reading in knowledge representation* (R. J. Brachman and H. J. Levesque, eds.) (Morgan Kaufman Publ., Los Altos), 41–70.
- [17] BRADLEY, S., HAX, A. and MAGNANTI, T., *Applied mathematical programming* (Addison Wesley, 1977).
- [18] BROOKE, A., KENDRICK, D. MEERAUS, A., *GAMS A User's Guide* (The Scientific Press, 1988).
- [19] BURGER, W. F., *MILD: a language and data base for modeling* (IBM Research Division, San Jose, Research Report RC 9639 (# 42311)).
- [20] CHIANG KAO, „A nonlinear programming expert system”, *Graduate School of Industrial Management* (National Cheng Kung University, Tainan, Taiwan, 1988).
- [21] CSÁKI P., MAYER J. és SZELKE E., „Döntési eljárások egy módszertani taxonómiai rendszere”, *Problémafeltáró tanulmány* (OMFB-EKPI, Budapest, 1988).
- [22] CUNNINGHAM, K. and SCHRAGE, L., *Optimizing Lotus 123 with VINO* (The Scientific Press, Redwood City, 1987).
- [23] DOLK, D. R. and KONSZYNSKI, B. R., „Knowledge representation for model management systems”, *IEEE Transactions on Software Engineering* Se-10 (1984), 619–628.
- [24] DOLK, D. R. and KONSZYNSKI, B. R., „Model management in organizations”, *Information & Management* 9 (1985), 35–47.
- [25] DOLK, D. R., „A generalized model management system for mathematical programming”, *ACM Transactions on Mathematical Software* 12 (1986), 92–126.
- [26] DOLK, D. R., „Model management and structured modeling: The role of an information resource dictionary system”, *Communications of the ACM* 31 (1988a), 704–718.
- [27] DOLK, D. R., „Model management systems for operations research a prospectus”, in *Mathematical methods for decision support* (G. Mitra, ed.) (Springer, 1988b), 347–373.
- [28] EDELMAYER A., „Objektum orientált programozási paradigmák”, *OTKA Tanulmány* (MTA SZTAKI, 1988).
- [29] FIKES, R. and KEHLER, T., „The role of frame-based representation in reasoning”, *Communications of the ACM* 28 (1985), 904–920.
- [30] FININ, T. and SILVERMAN, D., „Interactive classification as a knowledge acquisition tool”, in *Expert database systems* (Kershner, ed.) (The Benjamin-Cummings Publ. Co., 1986), 79–90.
- [31] FOURER, R., „Modeling languages versus matrix generators for linear programming”, *ACM Transactions on Mathematical Software* 9 (1983), 143–183.
- [32] FOURER, R., GAY, D. M. and KERNIGHAN, B. W., „AMPL: a mathematical programming language”, *Technical Report 87-03* (Northwestern University, Evanston III., 1987).
- [33] FOURER, R., GAY, D. M. and KERNIGHAN, B. W., „A modeling language for mathematical programming”, *Management Science* 36 (1990), 519–554.
- [34] GALAMBOS, J. A., ABELSON, R. P. and BLACK, J. B., *Knowledge structures* (Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, London, 1986).
- [35] GENERAL OPTIMIZATION INC., *What's Best!* (Holden-Day, 1986).
- [36] GEOFFRION, A. M., „Modeling approaches and systems related to structural modeling”, *Working Paper No. 339* (Western Management Science Institute, Univ. of California, Los Angeles, 1987a).
- [37] GEOFFRION, A. M., „An introduction to structured modeling”, *Management Science* 33 (1987b), 547–588.
- [38] GEOFFRION, A. M., *Computer-based modeling environments: a road to greater productivity, quality and popularity for management science operations research, keynote address at the 1988 Canadian Operations Research Society Meeting in Montreal* (Univ. of California, 1988).

- [39] GLOVER, F., JONES, G., KARNEY, D., KLINGMAN, D., and MOTE, J., „An integrated production, distribution and inventory planning system”, *Interfaces* 9 (1979), 21–35.
- [40] GLOVER, F. and KLINGMAN, D., „Mathematical optimization — a successful tool for logistics problems”, *Operational Research* (J. Brans, ed.) (North-Holland Publishing Co., 1981).
- [41] GREENBERG, H. J., „A functional description of ANALYZE: a computer-assisted analysis system for linear programming models”, *ACM Transactions on Mathematical Software* 9 (1983), 19–56.
- [42] HARTMAN, D., „Selection and evaluation of structural optimization strategies by means of expert systems”, in *Expert systems in structural analysis* (A. Niku-Lari, ed.) (Pergamon Press, 1986), 39–56.
- [43] HAYES-ROTH, F., „Rule-based systems”, *Communications of the ACM* 28 (1985), 921–932.
- [44] HELFERICH, O. K., „An introduction to logistics decision support system”, *Computers in Distribution* (Auerbach Publishers Inc., 1983).
- [45] HERNÁDI, Á., BODÓ Z. és KNUTH, E., „A tudásábrázolás technikai és gépi eszközei”, *MTA SZTAKI Tanulmányok* (Budapest, 197/1987).
- [46] HERNÁDI, Á., „Új szközök a fogalmi modellezésben”, *MTA SZTAKI Tanulmányok* (Budapest, 205/1988).
- [47] HO, J. K., „Nonprocedural implementation of mathematical programming algorithms”, in *Proceedings of the ORSA Symposium on Recent Advances in Computers in Operations Research* (1988).
- [48] HOLSAPPLE, C. W. and WHINSTON, A. B., *Decision support systems: theory and applications* (Springer, 1987).
- [49] HÜRLIMANN, T. and KOHLAS, J., „LPL: a structured language for linear programming modeling”, *OR Spektrum* 10 (1988), 55–63.
- [50] JACKSON, P., *Introduction to expert systems* (Addison-Wesley Publishing Company, 1990).
- [51] KALL, P., *Stochastic linear programming* (Springer, 1976).
- [52] KALL, P. and MAYER, J., *SLP-IOR: a model management system for stochastic linear programming — system design —*, Manuscripte (Institut für Operations Research der Universität Zürich, 1991).
- [53] KATZ, S., RISMAN, L. J. and RODEH, M., „A system for constructing linear programming models”, *IBM Systems Journal* 19 (1980), 505–520.
- [54] KENDRICK, D., MEERAUS, A. and ALATORRE, J., *The planning of investment programs in the steel industry* (The John Hopkins University Press, 1983).
- [55] KLAHR, P. and WATERMAN, D. A., eds, *Expert systems techniques, tools and application* (Addison-Wesley Publ. Co., 1986).
- [56] KLINGMAN, D., PHILLIPS, N., STEIGER, D., WIRTH, R., PADMAN, R. and KRISHNAN, R., „An optimization based integrated short-term refined petroleum product planning system”, *Management Science* 33 (1987), 813–830.
- [57] KOCH PÉTER, „Tudásbázis elemzés”, *Információ Elektronika* 1–2 (1990), 66–77.
- [58] KRIBBE, W. and VAN DER HOEK, G., „Applicability of nonlinear programming in decision support systems: realization of user-friendly software”, *Report 8725/A* (Econometric Institute, Erasmus University, Rotterdam, 1988).
- [59] KRISHNAN, R., „PDM: a knowledge based tool for model construction”, in *Proceedings of the Twenty Second Annual Hawai Conference on the System Sciences* (Kona, HI, 1988).
- [60] KRISHNAN, R., „PM*: A logic modeling language for model construction”, *Decision Support Systems* (1988).
- [61] LENARD, M. L., „Structured model management”, in *Mathematical methods for decision support* (G. Mitra, ed.) (Springer, 1988), 375–391.
- [62] *Level Five Research, INSIGHT 2+*, Reference Manuel, Version 1.0 (Florida, 1986).
- [63] LIEBMAN, J., LASDON, L., SCHRAGE, L. and WAREN, A., *Modeling and optimization with GINO* (The Scientific Press, Palo Alto, 1986).

- [64] LOTUS Development Corporation, *LOTUS 1-2-3, User's Manual* (Cambridge, Massachusetts, 1983).
- [65] LUCAS, C. and MITRA, G., *CAMPS: preliminary user manual* (Department of Mathematics and Statistics, Brunel University, U.K., 1985).
- [66] MAROS, I., „A general phase-I method in linear programming”, *European Journal of Operations Research* 23 (1986), 64-77.
- [67] MAROS, I., *MILP mikrogépen futó lineáris programozási rendszer, Rövid felhasználói dokumentáció a V1.30 verzióhoz* (1988).
- [68] MAROS, I., „MILP linear programming optimizer for personal computers under DOS”, *Preprints in Optimization* (University of Braunschweig, 1990).
- [67] MARSTEN, R. E., „The design of the XMP linear library”, *ACM Transactions on Mathematical Software* 7 (1981), 481-497.
- [68] MATURANA, V., *Comparative analysis of mathematical programming systems* (Western Management Science Institute, University of California, Los Angeles, 1987).
- [69] MAYER, J., „Probabilistic constrained programming: a reduced gradient algorithm implemented on PC”, *Working Paper-88-39* (International Institute for Applied Systems Analysis, 1988).
- [70] MCALLISTER, P. H., STONE, J. C., DANTZIG, G. B., „An interactive model management system: user interface and system design”, *Systems Optimization Laboratory, Stanford University, SOL 90-3* (1990).
- [71] MCCORMICK, G. P., *Nonlinear programming* (John Wiley & Sons, 1983).
- [72] MINSKY, M., „A framework for representing knowledge”, in *The psychology of computer vision* (P. Winston, McGraw-Hill, ed.); in *Mind and design* (J. Haugeland, ed.) (The MIT Press, Cambridge, 1981), 95-128.
- [73] MITRA, G., „Models for decision making: an overview of problems tools and major issues”, in *Mathematical methods for decision support* (G. Mitra, ed.) (Springer, 1988), 17-53.
- [74] MURPHY, F. H. and STOHR, E. A., „An intelligent system for formulating linear programs”, *Decision Support Systems* 2 (1986), 39-47.
- [75] MURTAGH, B. A. and SAUNDERS, M. A., *MINOS User Guide*, Technical Report 77-9 (Stanford, 1977).
- [76] MURTAGH, B. A. and SAUNDERS, M. A., „Large scale linearly constrained optimization”, *Mathematical Programming* 14 (1978), 41-72.
- [77] MURTAGH, B. A. and SAUNDERS, M. A., „A projected Lagrangian algorithm and its implementation for sparse nonlinear constraints”, *Mathematical Programming Study* 16 (1982), 84-117.
- [78] MURTAGH, B. A. and SAUNDERS, M. A., *MINOS 5.0 User Guide*, Technical Report Sol 83-20 (Stanford, 1983).
- [79] NANCE, R. E., BALCI, O. and BARGER, L. F., *Specification languages: understanding their role in simulation model development*, Technical Report, TR-87-7 (Department of Computer Science, Virginia Polytechnic).
- [80] NEMHAUSER, G. L. and WOLSEY, L. A., *Integer and combinatorial optimization* (John Wiley & Sons, 1989).
- [81] ÖREN, T. I., ELZAS, M. S. and ZEIGLER, B. P., *Theory of modeling and simulation* (John Wiley & Sons, 1984).
- [82] PALMER, K., *A model management framework for mathematical programming* (John Wiley & Sons, 1984).
- [83] PAPADIMITRIU, C. H. and STEIGLITZ, K., *Combinatorial optimization: algorithms and complexity* (Prentice-Hall, 1982).
- [84] PAPALAMBROS, P., „Knowledge-based systems in optimal design”, in *Computer aided optimal design: structural and mechanical systems* (Mota Soares, ed.) (Springer, 1987).
- [85] PÁSZTOR, Z. és SÁNTÁNÉ- TÓTH, E., „Számítógépes szakértő rendszerek alkalmazásának és tervezésének kérdései”, *Információ és Elektronika* 4-5 (1987), 153-158.

- [86] POWELL, P., „Decision support systems, the user and the language of OR”, Preprint.
- [87] RAPCSÁK, T., „Az operációkutatás kialakulásáról és hazai helyzetéről”, *Magyar Tudomány* 4 (1988), 259–265.
- [88] RAPCSÁK, T., „Heurisztikákat és mesterséges intelligencia módszereket is alkalmazó matematikai programozási eszközszoftver”, *Tanulmány* (OMFB-EKPI, Budapest, 1989).
- [89] RAYNA, G., *REDUCE: software for algebraic computation* (Springer, 1987).
- [90] ROY, A., LASDON, L. and LORDEMAN, J., „Extending planning languages to include optimization capabilities”, *Management Science* 32 (1986), 360–373.
- [91] ROY, A., LASDON, L. and PLANE, D., „End-user optimization with spreadsheet models”, *European Journal of Operational Research* 39 (1989), 131–137.
- [92] RUSZCZYŃSKI, A., „A regularized decomposition method for minimizing a sum of polyhedral functions”, *Mathematical Programming* 35 (1986), 309–333.
- [93] SÁNTÁNÉ TÓTH EDIT, „Az ismeretalapú rendszerek építésének módszertani kérdései”, *Információ Elektronika* 1-2 (1990), 77–97.
- [94] SCHITTKOWSKI, K., *EMP: an expert system for mathematical programming* (Mathematisches Institut, Universität Bayreuth, 1987).
- [95] SCHRAGE, L., *Linear, integer and quadratic programming with LINDO* (The Scientific Press, 1984).
- [96] SCHRAGE, L., CUNNINGHAM, K., *Demo LINGO/PC: Language for Interactive general Optimization, version 1.04a* (LINDO System inc., Chicago, 1988).
- [97] SCHULZE, K. and CRYER, C. W., *NAXPERT: a prototyp expert system for numerical software* (Institut für Numerische und instrumentelle Mathematik Westfälische Wilhelms-Universität zu Münster, 1987).
- [98] SHARDA, R., „Optimization using spreadsheets on a microcomputer”, *Annals of Operations Research* 5 (1986), 599–612.
- [99] SIMONS, R., „Mathematical programming modeling using MGG”, *IMA Journal of Mathematics in Management* 1 (1987), 267–276.
- [100] SOÓS, Zs., *NEQP*, MTA SZTAKI Working Paper MO/63 (Budapest, 1986).
- [101] SOÓS, Zs., *RQP/NLP*, MTA SZTAKI Working Paper MO/71 (Budapest, 1986).
- [102] SOWIZRAL, H. A. and KIPPS, J. R., *ROSIE: a programming environment for expert systems*, R-3246-ARPA (The Rand Corporation, July, 1985).
- [103] SPRAGUE, R. H. and CARLSON, E. D., *Building effective decision support systems* (Prentice Hall, 1982).
- [104] STRAZICKY, B., „TWO-STAGE: a code of a basis decomposition method for stochastic programming”, *IIASA Working Paper*, WP-87-126 (1987).
- [105] *Szakértői rendszerek '88', Ismeretalapú információfeldolgozás Magyarországon*, Szerk: Gábor András (Számítástechnika-Alk. Vállalat, Budapest, 1988).
- [106] SZÁNTAI, T. and EDWARDS, J., „PCSP user's manual”, in Documentation for the ADO/SDS collection of stochastic programming codes (Edwards, ed.), IIASA Working Paper, WP-85-02 (1985).
- [107] SZÁNTAI, T., „Evaluation of a special multivariate gamma distribution function”, *Mathematical Programming Study* 27 (1986), 1–16.
- [108] SZÁNTAI, T., „A computer code for solution of probabilistic-constrained stochastic programming problems”, in Numerical techniques for stochastic optimization (Yu. Ermoliev and R. J.-B. Wets, eds.) (Springer-Verlag, Berlin, Heidelberg, New York, 1988), 229–235.
- [109] SZELKE, E., „Piaci felmérés a döntési módszerek programcsomagjairól”, *Döntéselőkészítő tanulmány* (OMFB-EKPI, Budapest, 1988).
- [110] TALUKDAR, S. N., CARDOZO, E., LEÃO, L., BANARES, R. and JOBBANI, R., „A system for distributed problem solving”, in Coupling symbolic and numerical computing in expert systems (J. S. Kowalik, ed.) (Elsevier Science Publ., 1986).
- [111] THIRIEZ, H., „OR Software”, *European Journal of Operational Research* 39 (1989), 223–224.

- [112] TURBAN, E. and WATKINS, P. R., „Integrating expert systems and decision support systems”, in *Decision support systems: putting theory into practice* (Prentice-Hall Inc., 1986).
- [113] VÁMOS, T., „Convergences of cooperative control, decision and expert systems”, *Proceedings of 25th Conference on Decision and Control* (Athens, Greece., December 1989), 616–621.
- [114] WELCH, J. S., „PAM-Practitioner's approach to modeling”, *Management Science* **33** (1987), 610–625.
- [115] ZELENY, M., *Multiple criteria decision making* (McGraw-Hill, 1982).

(Beérkezett: 1991. június 28.)

(Átdolgozva beérkezett: 1992. október 26.)

BÍRÓ MIKLÓS, MAYER JÁNOS, RAPCSÁK TAMÁS ÉS VERMES MÁTYÁS
MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1111 BUDAPEST, KENDE U. 13–17.

ON EXPERT SYSTEMS OF MATHEMATICAL PROGRAMMING

M. BÍRÓ, J. MEYER, T. RAPCSÁK AND M. VERMES

The paper investigated the requirements, which ensure that a mathematical programming modeling system or a model management system becomes a mathematical programming expert system, i.e. an intelligent software tool for solving mathematical programming problems or decision problems containing optimization problems.

A mathematical programming model management system has been developed in the Operations Research Department of the Computer and Automation Institute of the Hungarian Academy of Sciences. The modeling language of the system is GAMS and the system includes a taxonomical classification of decision models.

BÁZISINVERZ NAGYMÉRETŰ, RITKÁS LINEÁRIS PROGRAMOZÁS FELADATOKBAN*

MÉSZÁROS CSABA

Budapest

A dolgozat rövid áttekintést ad a nagyméretű, ritkás lineáris programozási (LP) feladatok gyakorlatban használt bázisinvertálási és bázisinverz újraszámítási módszereiről. A különböző invertáló eljárásokat gyakorlati feladatokból származó, numerikusan rosszul viselkedő bázisokkal teszteltük. Egy új pivotválasztó stratégiát is bemutatunk, mely az invertálás során figyelembe veszi a korábbi transzformációk numerikus hatásait. Az új pivotválasztó stratégiától kedvezőbb numerikus viselkedést vártunk, amit az elvégzett kísérletek tapasztalatai alá is támasztottak.

1. Bevezetés

A lineáris programozás tudományos, mérnöki, közgazdasági felhasználásában gyakran találkozhatunk nagyméretű LP feladatokkal. Az ilyen feladatok nemritkán több ezer feltételből és változóból is állhatnak, és jellemzőjük, hogy a feltételi mátrix elemei között nagyon kevés a nullától különböző elem. Ezres nagyságrendnél a feltételi mátrix kitöltöttségére (azaz a nemnulla elemek számának arányára) egy százalék a tipikus érték. Ezt a tulajdonságot ritkasságnak nevezik. Az ilyen LP feladatok megoldása sok alapvető kérdést vetett fel a matematikában és a számítástudományban. A lineáris programozás ritkasságot kihasználó számítógépes implementációi az aktuális bázismátrix periódikus újrainvertálását végzik az iterációk műveletigényének csökkentése, és az ebből adódó kedvező numerikus hatások miatt. A megbízhatóan működő invertáló rutin az egyik legfontosabb tartozéka az LP programcsomagoknak. Az invertáló algoritmusokkal szemben támasztott négy legfontosabb igény a következő:

1. Megbízható működés (numerikus stabilitás)
2. Rövid éta file generálása
3. Gyors futási idő
4. Az iterációk során gyors FTRAN és BTRAN műveletek végzése

A dolgozat célja a leggyakrabban használatos invertáló eljárások és LP rendszerekbe illesztésük vizsgálata az előző szempontok alapján.

A második fejezet az inverz „hagyományos” szorzatformájával (Product From of Inverse, PFI), a harmadik fejezet az eliminációs formával (Elimination From of Inverse, EFI) és a dekompozíció báziscsere utáni fenntartásával foglalkozik. A

*A dolgozat a 2587 számú OTKA szerződés részbeni támogatásával készült.

negyedik fejezet az invertálás pivotválasztó algoritmusai közül mutat be néhány fontosabbat. Az ötödik fejezet implementációs problémákat és számítógépes kísérleti eredményeket vizsgál.

Ezúton szeretnék köszönetet nyilvánítani Maros Istvánnak a dolgozat elkészítéséhez nyújtott jelentős szakmai segítségéért, és a számítógépes kísérleteim „alanyául” szolgáló mintegy 20, numerikusan rosszul viselkedő LP bázis összegyűjtéséért és rendelkezésemre bocsátásáért.

2. A bázisinverz szorzatformája (PFI)

A „Product Form of Inverse” (PFI) típusú szimplex módszer a bázis inverzét a Gauss-Jordan elimináció alkalmazásával elemi transzformációmátrixok szorzataként állítja elő és tárolja [3,4,8]. Először nézzük meg egy új bázisra áttérés esetét :

Az m -ed rendű

$$B = (b_1, b_2, \dots, b_m)$$

bázisról áttérünk a

$$B_a = (b_1, b_2, \dots, b_{s-1}, a, b_{s+1}, \dots, b_m)$$

bázisra, vagyis a b_s oszlopot egy a oszlopra cseréljük. Ekkor B_a^{-1} a következőképpen származtatható B^{-1} -ből :

Legyen az a bázisbeli előállítás

$$a = \sum_{i=1}^m y_i b_i .$$

Nyilvánvaló, hogy B_a akkor és csak akkor bázis, ha $y_s \neq 0$, ezért

$$b_s = \frac{1}{y_s} a - \sum_{\substack{i=1 \\ i \neq s}}^m \frac{y_i}{y_s} b_i ,$$

azaz $b_s = B_a \eta$, ahol

$$\eta^T = \left(-\frac{y_1}{y_s}, \dots, -\frac{y_{s-1}}{y_s}, \frac{1}{y_s}, -\frac{y_{s+1}}{y_s}, \dots, -\frac{y_m}{y_s} \right) .$$

Az

$$F = \begin{bmatrix} 1 & & \eta_1 & & 0 \\ & \ddots & \vdots & & \\ & & \eta_s & & \\ & & \vdots & \ddots & \\ 0 & & \eta_m & & 1 \end{bmatrix}$$

jelöléssel $B = B_a F$, azaz $B_a^{-1} = F B^{-1}$. Ezt a gondolatmenetet használhatjuk az inverz előállítására. Kiindulva az egységmátrixból B oszlopait egyesével cseréljük ki I oszlopaival, minden lépésnél eltávolítva az F_i mátrixot. A továbbiakban feltételezzük, hogy

$$b_1, \dots, b_m$$

sorrendben az

$$e_1, \dots, e_m$$

helyére lép be. Így a következő előállítást kapjuk :

$$B^{-1} = F_m F_{m-1} \cdots F_1 F_0,$$

ahol $F_0 = I$,

$$F_i = (e_1, \dots, e_{i-1}, \eta_i, e_{i+1}, \dots, e_m),$$

$$\eta_i^T = \left(-\frac{y_{1i}}{y_{ii}}, \dots, -\frac{y_{i-1,i}}{y_{ii}}, \frac{1}{y_{ii}}, -\frac{y_{i+1,i}}{y_{ii}}, \dots, -\frac{y_{mi}}{y_{ii}} \right),$$

és

$$y_i = F_{i-1} F_{i-2} \cdots F_1 F_0 b_i.$$

Itt megjegyezzük, hogy mivel az F_i mátrixot az i index (a pivot index) és az η_i vektor egyértelműen meghatározza, ezért az F_i mátrixok tárolása helyett elegendő az η_i vektorokat tárolni. Mivel egy bázisban gyakran vannak logikai változók, (és ezeknek egységvektorok felelnek meg) a gyakorlatban ez azt jelenti, hogy a teljes szorzatformához legfeljebb m vektor tárolása szükséges.

A PFI típusú invertálás általános váza tehát a következő : Legyen I^s azon sorindexek, J^s azon oszlopindexek halmaza, melyen az s -edik lépésig nem pivotáltunk, B^s pedig a még be nem vont oszlopokból álló almatrix (ezt a későbbiekben **aktív almatrixnak** nevezzük), és jelölje b_{ij}^s B^s elemeit.

0. Inicializálás : $s = 0$, $B^0 = B$, $I^0 = \{1, \dots, m\}$, $J^0 = \{1, \dots, m\}$.
1. Keressünk egy (r, t) pivotpozíciót, úgy, hogy $b_{rt}^s \neq 0$ és $(r, t) \in I^s \times J^s$.
2. Legyen $I^{s+1} = I^s \setminus \{r\}$ és $J^{s+1} = J^s \setminus \{t\}$.
3. B^{s+1} számítása :

$$b_{ij}^{s+1} = b_{ij}^s - \frac{b_{rj}^s b_{it}^s}{b_{rt}^s}, \quad i = \{1 \dots m\} \setminus \{r\}, \quad j \in J^{s+1},$$

$$b_{rj}^{s+1} = \frac{b_{rj}^s}{b_{rt}^s}, \quad j \in J^{s+1}.$$

4. Tároljuk az

$$\eta_s = \left(-\frac{b_{1t}^s}{b_{rt}^s}, \dots, -\frac{b_{r-1,t}^s}{b_{rt}^s}, \frac{1}{b_{rt}^s}, -\frac{b_{r+1,t}^s}{b_{rt}^s}, \dots, -\frac{b_{mt}^s}{b_{rt}^s} \right)^T$$

vektort.

5. Legyen $s = s + 1$ és ha $J^s \neq \emptyset$ menjünk az 1. lépésre.

Az eljárás végén egy

$$B^{-1} = F_m P_m F_{m-1} P_{m-1} \cdots F_1 P_1 Q_1 \cdots Q_m$$

előállítást kapunk, ahol F_i elemi transzformációmátrix, P_i és Q_i pedig elemi permutációmátrixok.

3. A bázis inverz eliminációs formája (EFI)

Egy más típusú invertálás gondolatát Harry Markowitz vetette fel 1957-ben, [9] dolgozatában. A módszer lényege, hogy a B bázist a következőképpen bontjuk fel: $B = LU$, ahol L alsó háromszögmátrix, U pedig felső háromszögmátrix. Ekkor az L és U mátrixokat könnyen szorzatformára írhatjuk, és így a következő előállítást kapjuk:

$$B^{-1} = U_1^{-1} U_2^{-1} \cdots U_m^{-1} L_m^{-1} L_{m-1}^{-1} \cdots L_1^{-1},$$

ahol

$$U_k = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & u_{kk} & & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix}$$

és

$$L_k = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & l_{k+1,k} & \ddots & \\ & & \vdots & \ddots & \\ 0 & & l_{mk} & & 1 \end{bmatrix}.$$

Megemlíttük, hogy U_k^{-1} és L_k^{-1} eta vektorként tárolható és elemei: $\eta_{ik} = -u_{ik}/u_{kk}$, vagy $\eta_{ik} = -l_{ik}/l_{kk}$, és „forward” illetve „backward” eta file az elnevezésük.

Az LU dekompozíció Gauss-eliminációval nyerhető:

Legyen $B^0 = B$ és B^s elemeit jelöljük b_{ij}^s -vel.

Az első lépésben a b_{i1}^0 , $(i = 2, \dots, m)$ elemeket elimináljuk (feltételezzük, hogy $b_{1,1}^0 \neq 0$). Az elimináció a következő transzformációval végezhető, mátrixát jelöljük L_1^{-1} -el:

$$L_1^{-1} = \begin{bmatrix} 1 & & & 0 \\ -b_{2,1}^0/b_{1,1}^0 & 1 & & \\ \vdots & & \ddots & \\ -b_{m,1}^0/b_{1,1}^0 & & & 1 \end{bmatrix},$$

továbbá legyen $B^1 = L_1^{-1}B^0$.

A k -adik lépésben a b_{ik}^{k-1} , $(i = k+1, \dots, m)$ elemeket elimináljuk az

$$L_k^{-1} = \begin{bmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & -b_{k+1,k}^{k-1}/b_{kk}^{k-1} & \ddots & \\ & & \vdots & & \ddots \\ & & -b_{mk}^{k-1}/b_{kk}^{k-1} & & 1 \end{bmatrix}$$

transzformációval, és legyen $B^k = L_k^{-1}B^{k-1}$. Az m -edik lépés végén az $U = B^m$ felső háromszögmátrixot kapjuk.

Az EFI típusú invertálás általános váza tehát a következő:

Legyen I^s azon sorindexek, J^s azon oszlopindexek halmaza, melyen az s -edik lépésig nem pivotáltunk, B^s pedig az aktív almatrix, azaz a mátrix azon része, amit még nem faktorizáltunk. Itt most az aktív almatrix definíciója más, mint az előző fejezetben, ugyanis most a sorok közül is elhagyjuk azokat, amelyekben már pivotáltunk, azaz ez esetben az aktív almatrix négyzetes mátrix, ellentétben a PFI esettel. Jelentse b_{ij}^s B^s elemeit.

0. Inicializálás: $s = 0$, $B^0 = B$, $I^0 = \{1, \dots, m\}$, $J^0 = \{1, \dots, m\}$.
1. Keressünk egy (r, t) pivotpozíciót, úgy, hogy $b_{rt}^s \neq 0$ és $(r, t) \in I^s \times J^s$.
2. Legyen $I^{s+1} = I^s \setminus \{r\}$ és $J^{s+1} = J^s \setminus \{t\}$.
3. B^{s+1} számítása:

$$b_{ij}^{s+1} = b_{ij}^s - \frac{b_{rj}^s b_{it}^s}{b_{rt}^s}, \quad i \in I^{s+1}, \quad j \in J^{s+1}.$$

4. Tároljuk az

$$\eta_s = \left(-\frac{b_{r+1,t}^s}{b_{rt}^s}, \dots, -\frac{b_{mt}^s}{b_{rt}^s} \right)^T$$

vektort.

5. Legyen $s = s + 1$, és ha $J^s \neq \emptyset$ menjünk az 1. lépésre.

Az eljárás végén egy

$$B^m = L_{m-1}^{-1} P_{m-1} \cdots L_1^{-1} P_1 B Q_1 \cdots Q_{m-1}$$

felbontást kapunk, ahol P_i és Q_i elemi permutációmátrixok, $U = B^m$ pedig felső háromszögmátrix.

Az új dekompozíció számítása báziscsere után

A szimplex módszer iterációs lépései közben báziscserére is sor kerül, azaz a bázis egy vektorát egy másikra cseréljük. A régi B és az új \hat{B} bázisok így csak egyetlen vektorban térnek el egymástól. Az ilyen típusú bázispárokat **szomszédos bázisoknak** nevezzük, és a szimplex módszer iterációk során szomszédos bázisok láncán keresztül jut el valamely megállási kritériumig. Ezért egy olyan gyors és numerikusan stabil eljárásra is szükségünk van, mely a régi bázis inverzéből az új bázis inverzét származtatja. Markowitz [9] dolgozatában erre még a 2. fejezetben leírt „hagyományos” elgondolást javasolja, vagyis a belépő vektor bázisbeli képéből elkészített újabb F_j elemi transzformációmátrix hozzávételét a meglévő felbontáshoz. Egy új módszert dolgozott ki Bartels és Golub, majd Bartels 1971-ben elméletileg is bizonyította [2] dolgozatában módszerének lényegesen jobb numerikus tulajdonságait Markowitz régi ötletével szemben. Ezt Saunders [14] ötletének felhasználásával J.K. Reid módosította 1975-ben [13]. Egy másik, ettől lényegesen különböző eljárást dolgozott ki Forrest és Tomlin 1971-ben, és kísérletileg kimutatták, hogy technikájuk a nemnulla elemek növekedési aránya szempontjából sokkal jobb, mint a Markowitz-féle implementáció [6].

Bartels és Golub módszere

Induljunk ki a 3. fejezet eredményeként adódott

$$U = L_{m-1}^{-1} P_{m-1} \cdots L_1^{-1} P_1 B Q$$

felbontásból. Az iteráció utáni \hat{B} új bázis egyetlen oszlopban tér el B -től, így fennáll a következő egyenlőség :

$$L_{m-1}^{-1} P_{m-1} \cdots L_1^{-1} P_1 \hat{B} Q = S ,$$

ahol S abban az egyetlen oszlopban tér el U -tól, amelyikben B és \hat{B} is különbözik egymástól, azaz S a következő alakú : (a * a mátrixban a lehetséges

nemnulla elemeket jelöli)

$$S = \begin{bmatrix} * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & * & * & * & * \\ & & * & & * & * \\ & & * & & & * \end{bmatrix}.$$

Ez az ún. csúcsos S mátrix. Ezután a „csúcsot” oszlopcserevel a leghátsó helyre vihetjük (a közbeeső oszlopokat eggyel balra léptetve), és így a következő, ún. felső Hessenberg mátrixot kapjuk :

$$H = \begin{bmatrix} * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix}.$$

Ezután az aldiagonálisban levő elemeket elimináljuk, a jobb numerikus viselkedéshez sorcserekkkel azt is elérhetjük, hogy mindig a h_{jj} , $h_{j+1,j}$ elemek közül a nagyobb abszolút értékűn pivotáljunk (Wilkinson parciális pivot stratégiája [16]). Jelölje ezen eliminációs mátrixokat

$$L_m^{-1}, L_{m+1}^{-1}, \dots, L_{m+n-1}^{-1},$$

a pivotálásokhoz tartozó permutációmátrixokat pedig

$$P_m^{-1}, P_{m+1}^{-1}, \dots, P_{m+n-1}^{-1}$$

(feltételezve, hogy n pivotálásra van szükség). Ekkor az új LU faktorizáció a következőképpen írható fel [1,7] :

$$\hat{U} = L_{m+n-1}^{-1} P_{m+n-1} \dots L_1^{-1} P_1 \hat{B} \hat{Q}.$$

Ez láthatóan egy numerikusan stabil eljárás, viszont a keletkező nemnulla elemek szempontjából nem túl ideális, ugyanis ha a „csúcs” oszlopa a p oszlop, akkor potenciálisan a p -től m -ig minden sorban és oszlopban az elemeket transzformálni kell, és ez az U mátrix viszonylag gyors feltöltődésével jár.

Saunders [14] ötlete, hogy a H mátrix ritkasságát kihasználva, a mátrix trianguláris részeinek leválasztásával a feltöltődés gyorsaságát csökkenteni lehet, amit egyszerű sor- és oszlopcserekkel lehet megtenni. Erre egy példa a következő,

az eredeti H mátrixból sor- és oszloppermutációkkal a \hat{H} mátrixot kapjuk :

$$H = \begin{bmatrix} * & & * & & \\ * & * & * & & \\ & * & & * & \\ & & * & * & * \\ & & & * & * \\ & & & & * & * \end{bmatrix} \rightarrow \hat{H} = \begin{bmatrix} * & * & & & * \\ & * & * & & \\ & & * & * & \\ & & & * & * \\ & & & * & * & * \\ & & & & * & * \end{bmatrix},$$

amelynél egyetlen pivotálás elegendő, míg az eredeti mátrixnál hat helyen lett volna szükség pivotálásra. Azaz Saunders ötletét alkalmazva (a triangularitás mértékétől függően) már csak egy kisebb blokkban kell elemeket eliminálni, és ez kisebb feltöltődést eredményez. A trianguláris részek leválasztásának algoritmusára megegyezik az invertálásnál szokásos eljárással, részletesebb ismertetése a 4. fejezetben található.

Forrest és Tomlin módszere

Tegyük fel, hogy a B bázis p -edik oszlopának helyére egy a vektor lép be, és \hat{B} az új bázismátrix, továbbá hogy a régi dekompozíció : $B = LU$.

Legyen $\alpha = L^{-1}a$, amit a bázisba lépő vektor **parciális transzformáltjának** nevezünk. Az előző pontban leírtak szerint $S = L^{-1}\hat{B}$ csak a p -edik oszlopában különbözik U -tól. Legyen $H = L^{-1}\hat{B}Q$ a felső Hessenberg mátrix (Q az a permutáció, amely az α oszlopot az m -edik pozícióba viszi, és $(p+1)$ -től m -ig az oszlopokat eggyel balra lépteti). Ezután H -ban a p -edik sorban a p -től $(m-1)$ -ig levő oszlopokban álló elemeket elimináljuk (azaz $h_{pp}, \dots, h_{p,m-1}$ elemeket, melyek az $u_{p,p+1}, \dots, u_{p,m}$ elemekkel megegyeznek), és így egy sorpermutált trianguláris mátrixot kapunk. Ezt az eliminációt az R^{-1} sortranszformáció fogja végrehajtani. Legyen R a következő alakú mátrix :

$$R = \begin{bmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ 0 & \dots & 0 & 1 & r_{p+1} & \dots & r_m \\ & & & & \ddots & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{bmatrix},$$

azaz $R = I + e_p r$, ahol e_p a p -edik egységvektor, és $r = (0, \dots, 0, r_{p+1}, \dots, r_m)$. Könnyen látható, hogy $R^{-1} = I - e_p r$.

Legyen

$$\hat{u} = (0, \dots, 0, u_{p,p+1}, \dots, u_{p,m})$$

és

$$\mathbf{r} = \hat{\mathbf{u}}U^{-1}.$$

Ekkor

$$\begin{aligned} R^{-1}U &= (I - \mathbf{e}_p \mathbf{r})U \\ &= U - \mathbf{e}_p \hat{\mathbf{u}}U^{-1}U \\ &= U - \mathbf{e}_p \hat{\mathbf{u}} \end{aligned}$$

és a legutolsó kifejezés az U mátrix olyan módosulása, melynél az $u_{p,p+1}, \dots, u_{pm}$ elemek nullák [3]. Most legyen $V = R^{-1}H$, ekkor V mátrix H -tól a

$$h_{pp}, \dots, h_{p,m-1}$$

elemekben különbözik (melyek V -ben nullák), és az utolsó oszlop $R^{-1}\alpha$. Nyilvánvaló, hogy V egy felső háromszögmátrix sorpermutációja, és $Q^{-1}V$ az \hat{U} felső trianguláris mátrix.

Az új felbontás a következőképpen írható:

$$\hat{B}^{-1} = QV^{-1}R^{-1}L^{-1},$$

vagy másképpen, felhasználva a $V = Q\hat{U}$ definíciót,

$$\hat{B}^{-1} = Q\hat{U}^{-1}Q^{-1}R^{-1}L^{-1}.$$

A V^{-1} szorzatformája nagyon könnyen származtatható U^{-1} szorzatformájából: az U_p^{-1} transzformációt töröljük, és minden elemet, mely U p -edik sorának felel meg, nullára állítunk. Az új transzformáció $R^{-1}\alpha$ -ból származik, és balról U^{-1} -hez írjuk. Érdemes észrevenni, hogy az \mathbf{r} vektor elkészítésénél a

$$\mathbf{z}U_k^{-1} = (0, \dots, 0, z_{p+1}, \dots, z_m)U_k^{-1}$$

műveleteknél csak a z_{p+1}, \dots, z_k elemek transzformálódnak, és a z_{k+1}, \dots, z_m elemek változatlanok maradnak, ugyanis U_k^{-1} felső háromszög elemi transzformáció, és $k > p$.

Ezek alapján az \mathbf{r} vektor számítása és U^{-1} módosítása a következőképpen végezhető el:

0. Inicializálás : Az \mathbf{r} sorvektor elemeinek kezdőértéke legyen nulla, $U_1^{-1}, \dots, U_{p-1}^{-1}$ transzformációkat átugorjuk, U_p^{-1} -t töröljük, és legyen $k = p + 1$.
1. Általános lépés : ha $u_{pk} \neq 0$, rakjuk ezt az elemet a sorvektor k -edik pozíciójába, és töröljük a neki megfelelő elemet U_k^{-1} -ből. Transzformáljuk a sorvektort U_k^{-1} -val.
2. Az 1. lépést $k = p + 1, \dots, m$ -re hajtsuk végre.

Az eljárás végén a sorvektorban kapjuk az r vektort. Ezután az $R^{-1}\alpha$ kiszámítása következik, majd balról a „backward” eta file-hoz írása. Az utolsó lépés R^{-1} jobbról a „forward” eta file-hoz adása. Ezzel elkészül az új bázis dekompozíciója.

A módszer közben lehetőség van egy újrainvertálási kritérium ellenőrzésére: Kihaszználjuk, hogy az új bázis determinánsa a régi bázis determinánsának és a pivotelem reciprokának szorzata, azaz

$$\det(\hat{B}^{-1}) = \frac{1}{\alpha_p} \det(B^{-1}) .$$

Forrest és Tomlin módszerében U_p^{-1} -t töröljük, és két új eta vektor lép be, melyekben a pivotelemek 1 és $\hat{\alpha}_p$, ahol az utóbbi az $R^{-1}\alpha$ vektor p -edik pozíciójában álló elem. Ezért

$$\det(\hat{B}^{-1}) = \left(\frac{u_{pp}}{\hat{\alpha}_p} \right) \det(B^{-1}) ,$$

azaz fennáll a következő egyenlőség:

$$\hat{\alpha}_p = \alpha_p u_{pp} .$$

Az $\hat{\alpha}_p$ érték tehát az $R^{-1}\alpha$ számítása közben ellenőrizhető, és az újrainvertálási kritérium az

$$\hat{\alpha}_p = \alpha_p u_{pp}$$

egyenlőség nem megfelelő teljesülése [6].

Forrest és Tomlin kísérleti eredményei azt bizonyították, hogy módszerükkel a feltöltődés üteme lényegesen kisebb akár Bartels és Golub, akár Markowitz algoritmusához viszonyítva. További javulást eredményezhet Saunders ötletének itteni alkalmazása, azaz a H mátrix trianguláris részeinek permutációkkal történő leválasztása, melynek eredményeként a szükséges transzformációkat a H mátrix egy kisebb blokkjában kell majd elvégezni.

4. Pivotválasztási algoritmusok

Az invertáló algoritmusokban a legfontosabb szerep a pivotelem-kiválasztó stratégiáknak jut. Egy rosszul választott pivotelem numerikus gondokat okozhat, és az eredetileg ritkás bázis jelentős feltöltődését is eredményezheti. Például, ha a

$$B = \begin{pmatrix} * & * & * & * & * \\ * & & & * & \\ * & & * & & \\ * & * & & & \\ * & & & & \end{pmatrix}$$

mátrix $b_{1,1}$ elemén pivotálunk, a következő lépésben (a pivotálás elvégzése után) egy teljesen kitöltött mátrixhoz jutunk, ha pedig a $b_{1,5}$, vagy $b_{5,1}$ pivotelemet használjuk, akkor nem keletkezik újabb nemnulla elem.

Látható, hogy az előző fejezetekben leírt

$$b_{ij}^{s+1} = b_{ij}^s - \frac{b_{rj}^s b_{it}^s}{b_{rt}^s}$$

típusú műveleteknél az okozhat numerikus problémákat, ha a b_{rt}^s pivotelem „túl kicsi”. Ezt a veszélyt megpróbáljuk elkerülni: Csak azokat a b_{ij}^s elemeket fogadjuk el pivotelemként, melyekre teljesül a következő, ún. **stabilitási kritérium**:

$$|b_{ij}^s| \geq u \cdot \max\{|b_{ik}^s|, \quad k \in J^s\}$$

EFI típusú invertálásnál, illetve

$$|b_{ij}^s| \geq u \cdot \max\{|b_{kj}^s|, \quad k = 1, \dots, m\}$$

PFI típusú invertálásnál, ahol u egy rögzített $0 < u \leq 1$ valós szám. Ekkor a mátrix elemeinek növekedése legfeljebb $(1 + \frac{1}{u})$ -szeres lehet, azaz, ha a kiindulási mátrix legnagyobb abszolút értékű elemét a -val, az invertálás során az aktív almatrixokban keletkezett legnagyobb abszolút értékű elemet d -vel jelöljük, úgy d -re a következő felső becslést adhatjuk:

$$d \leq a \cdot \left(1 + \frac{1}{u}\right)^{m-1}.$$

Reid eredménye [13], hogy ez esetben az LU faktorizációnál keletkező E perturbációs mátrix elemeit a következőképpen becsülhetjük:

$$|e_{ij}| \leq 3.01 \cdot d \cdot m \cdot g,$$

ahol g a lebegőpontos műveletek végzésének gépi pontossága, E pedig az $LU = B + E$ perturbációs mátrix. A PFI típusú algoritmusnál előfordulhat, hogy nem találunk az u toleranciának megfelelő pivotelemet. Ekkor a szokásos technika, hogy u értékét ideiglenesen lecsökkentjük (általában 100-ad, vagy 1000-ed részére), és ennek a toleranciának megfelelő pivotelemet keresünk, majd u -t visszaállítjuk az eredeti értékére. Ezt nevezik **második fázisú pivotkeresésnek**. Gyakori módszer a stabilitás ellenőrzésére, hogy az invertálás közben állandóan figyeljük az aktív almatrixban a legnagyobb abszolút értékű elemet, és ha a kiindulási mátrixban levő legnagyobb abszolút értékű elemhez viszonyítva túlságosan nagy lesz (a szokásos kritérium $10^7 - 10^{10}$ -szeres növekedés), akkor az algoritmust megszakítjuk és egy új u pivotválasztási toleranciával indítjuk újra [15]. Az új tolerancia szokásos definíciója: $u_{\text{új}} = \min(2 \cdot u_{\text{rég}}, 1.0)$. Ez a módszer használható akkor is, ha a PFI második fázisban sem talál pivotelemet [15].

Gyakran a bázismátrix jelentős része sor- és oszloppermutációkkal triangulázható. A trianguláris részek technikailag nagyon kedvezőek, mert pivotáláskor az aktuális almátrixot nem kell transzformálni. A bázismátrix trianguláris részeit a következőképpen lehet megkeresni:

XTRIAN algoritmus

0. Inicializálás

Vezessünk be sor- és oszlopszámlálókat, azaz számoljuk és tároljuk a nemnulla elemek számát minden sorban és oszlopban. Legyen a j sor sorszámlálója r_j , a j oszlop oszlopszámlálója c_j .

1. Oszlopszámlálók vizsgálata

Keressünk olyan oszlopot, melynek oszlopszámlálója 1. Ha nincs ilyen oszlop, menjünk a 2. lépésre.

Ha a t oszlopra $c_t = 1$, akkor keressük meg az oszlopban található nemnulla elemet, legyen ennek sorindexe r .

Készítsük és tároljuk el a t oszlopból az eta vektort.

EFI algoritmus esetében készítsük és tároljuk el az r sorból készített eta vektort, PFI algoritmus esetében pedig transzformáljuk az r sort az 2. fejezetben leírtak szerint.

Menjünk végig az r sor minden nemnulla elemén, és csökkentjük eggyel azon oszlopok oszlopszámlálóit, melyekhez tartoznak.

Zárjuk ki a t oszlopot és az r sort a további vizsgálatokból.

Menjünk vissza az 1. lépéshez.

2. Sorszámlálók vizsgálata

Keressünk olyan sort, melynek sorszámlálója 1. Ha nincs ilyen sor, menjünk a 3. lépésre.

Ha az r sorra $c_r = 1$, akkor keressük meg a sorban található nemnulla elemet, legyen ennek oszlopindexe t .

Készítsük, és tároljuk el a t oszlopból az eta vektort.

Menjünk végig a t oszlop minden nemnulla elemén, és csökkentjük eggyel azon sorok sorszámlálóit, melyekhez tartoznak.

Zárjuk ki a t oszlopot és az r sort a további vizsgálatokból.

Menjünk vissza a 2. lépéshez.

3. STOP

Az XTRIAN algoritmus végén az aktuális almátrixot, melyben már nem található 1 számlálójú sor és oszlop, bázismagnak nevezzük. A következő négy algoritmust a bázismagban való pivotkeresésnél használhatjuk.

Hellerman és Rarick módszere

Hellerman és Rarick 1971-ben megjelent dolgozata [8] a témakör egyik legkorábbi jelentősebb eredménye. Algoritmusuk, melynek az „előre meghatározott pivot

eljárás” nevet adták, a bázismag úgynevezett erőltetett triangularizációját végzi PFI típusú invertálásnál. Ez pontosabban azt jelenti, hogy a bázis oszlopait két részre osztjuk, sorra félreteszünk bizonyos oszlopokat (közben csökkentjük a másik részben a megfelelő sorszámlálókat) egészen addig, amíg egy olyan sor nem keletkezik, melynek sorszámlálója 1. Ekkor itt pivotálunk, és az új aktív almatrix számításakor csak az előzőleg félretett oszlopokat kell transzformálni. Ezután a félretett oszlopok azon soraiban kereshetünk pivotelemet, melyek sorszámlálója az oszlopok másik partíciójában 0. Nyilvánvalóan olyan oszlopokat kell félretennünk, melyek a „lehető legjobban” elősegítik a megmaradt rész triangulárisá tételét. Erre adott egy ötletes megoldást Hellerinan és Rarick, és bár a [8] dolgozatukban közölt algoritmus nem veszi figyelembe a stabilitási kritériumot, nem nehéz olyan módosítást adni, amely a kritériumot is ellenőrzi. Az így módosított algoritmus a következő:

Jelölések :

- C : pivot oszlopok sorozata;
- R : pivot sorok sorozata;
- I : oszlopszámlálók;
- J : sorszámlálók;
- S : a félretett oszlopok halmaza;
- L : a félretett oszlopok száma;
- $t_k(n)$: az n oszlop azon nemnulla elemeinek száma, amelyhez tartozó sorok sorszámlálója $\leq k$;
- μ : a szükséges pivotálások száma.

HELLERMAN-RARICK algoritmus

0. A bázis trianguláris részeinek megkeresése (XTRIAN).

Jelöljük C -hez tartozónak a trianguláris oszlopokat, R -hez tartozónak a trianguláris sorokat.

1. Inicializálás

Számítsuk ki az I oszlopszámlálókat és J sorszámlálókat. Legyen μ a bázismag rendje, $L = 0$, és $S = \emptyset$.

2. Befejezés vizsgálata

- a) Ha $\mu = 0$ menjünk a 11. lépésre.
- b) Ha $\mu > 0$, akkor
 - (i) legyen k minimális, nullánál nagyobb sorszámláló.
 - (ii) Defináljuk N -et azon vektorok halmazaként, melyek nem tartoznak C -be, vagy S -be.

3. Számítsuk ki $t_k(n)$ értékét minden $n \in N$ -re.

4. Oszlopválasztás

- a) Ha $\max t_k(n) > 1$, akkor
 - (i) ha csak egy ilyen oszlop van, jelöljük az oszlop indexét α -val, menjünk az 5. lépésre.
 - (ii) Ha több ilyen oszlop van, válasszunk közülük egy maximális oszlopszámlálójú oszlopot, jelöljük az indexét α -val, és menjünk az 5. lépésre.

- b) Ha $\max t_k(n) = 1$, akkor
 - (i) definiáljuk N -et azon vektorok halmazaként, melyekre $t_k(n) = 1$.
 - (ii) Definiáljuk k új értékét annak a legkisebb sorszámlálónak, mely k régi értékénél nagyobb.
 - (iii) Menjünk a 3. lépésre.
- 5. *Az o oszlop ideiglenes kizárása*
 - a) Legyen $L = L + 1$.
 - b) Csatoljuk az o oszlopot S -hez.
 - c) Csökkentsük minden olyan sor sorszámlálóját, melyekhez az o oszlop tartalmaz nemnulla elemet.
- 6. *Sorszámlálók vizsgálata*
 - a) Ha $\min J > 1$, menjünk a 2.b)(i) lépésre.
 - b) Ha $\min J = 1$, és csak egy ilyen sor van, vagy ha $L = 0$, akkor
 - (i) jelöljük ezen oszlopnak az indexét u -val.
 - (ii) Menjünk a 9. lépésre.
 - c) Ha $\min J = 1$, és több ilyen sor van, akkor
 - (i) legyen $k = 1$.
 - (ii) Legyen N azon vektorok halmaza, melyek nincsenek C -ben, vagy S -ben.
- 7. *Számítsuk ki $t_k(n)$ értékét minden $n \in N$ -re*
- 8. *Pivotválasztás*
 - a) Ha $\max t_k(n) > 1$, akkor
 - (i) ha csak egy ilyen oszlop van, jelöljük indexét u -val, menjünk a 9. lépésre.
 - (ii) Ha több ilyen oszlop van, válasszunk közülük egy maximális oszlopszámlálójú oszlopot és jelöljük indexét u -val, menjünk a 9. lépésre.
 - b) Ha $\max t_k(n) = 1$, akkor
 - (i) definiáljuk N -et azon vektorok halmazaként, melyekre $t_k(n) = 1$.
 - (ii) Definiáljuk k új értékét annak a legkisebb sorszámlálónak, mely k régi értékénél nagyobb.
 - (iii) Menjünk a 7. lépésre.
- 9. *Pivotelemvizsgálat*
 - a) Keressük meg a pivotsort, azaz az u oszlopban azt a sort, melynek sorszámlálója 1, és jelöljük t -vel.
 - b) Ha a talált pivotelem nem teljesíti a stabilitási kritériumot, akkor az u oszlopot csatoljuk S -hez.
 - (i) Legyen $o = u$.
 - (ii) Menjünk az 5. lépésre.
 - c) Ha teljesül a stabilitási kritérium, akkor
 - (i) csatoljuk u -t C -hez, és zárjuk ki az u oszlopot a további pivotálások közül.
 - (ii) Csatoljuk t -t R -hez, és zárjuk ki a t sort a további pivotálások közül.
 - (iii) Végezzük el az S -be rakott oszlopok transzformációját, készítsük és tároljuk el a megfelelő eta vektort.
 - (iv) Csökkentsük azon sorok sorszámlálóit, melyeknél az u oszlopban nemnulla elem található.

(v) Legyen $\mu = \mu - 1$, ha $\mu = 0$, akkor menjünk a 11. lépésre.

10. Pivotkeresés az ideiglenesen elhagyott oszlopok között

a) Az ideiglenesen elhagyott oszlopok közül keressünk olyat, aminek van olyan nemnulla eleme, amely sorának sorszámlálója 0 (az oszlopok másik partíciójában), és teljesül a stabilitási kritérium.

(i) Ha nincs ilyen, menjünk a 6. lépésre.

(ii) Ha van ilyen, csatoljuk a sort R -hez, az oszlopot C -hez, és zárjuk ki őket a további pivotálások közül. Legyen $L = L - 1$, és vegyük ki a pivotoszlopot S -ből. Végezzük el az S -be rakott oszlopok transzformációját, készítsük és tároljuk el a megfelelő eta vektort. Menjünk a 9.c)(v) lépésre.

11. STOP

Röviden összefoglalva, az algoritmus olyan oszlopot fog ideiglenesen elhagyni, amellyel a lehető legtöbb minimális sorszámlálójú sorban csökkenteni lehet a sorszámlálót, illetve több ilyen oszlop létezése esetén ezek közül a maximális oszlopszámlálójút. Ha ily módon csak egy sorban sikerülne a sorszámlálót csökkenteni, akkor azt az oszlopot hagyjuk el, amely a minimális és a minimálisnál nagyobb, legkisebb sorszámlálójú sorok közül a lehető legtöbbször csökkent, stb., majd ugyanígy választhatunk pivotoszlopot is. Ezt az elvet a $t_k(n)$ függvény ügyes alkalmazásával lehet gyorsan megvalósítani.

Az algoritmussal végzett kísérletek tanulsága az volt, hogy az algoritmus gyorsasága megfelelő, ami annak az eredménye, hogy csak az előre meghatározott oszlopokat kell egy lépésben transzformálnunk, a többi oszlopot ki sem kell másolni az LP feladat mátrixából egy külön munkaterületre, és ezzel nagyon sok memóriaműveletet is megtakaríthatunk. Az erőltetett triangularizációs elvnek köszönhetően viszonylag rövid eta file-t kapunk. Numerikusan azonban nagy problémák léptek fel. Ez azzal magyarázható, hogy az S halmazbeli oszlopok közül csak meg lehetőséssel ritkán választhatunk pivotot, azaz ha egy oszlop bekerül az ideiglenesen elhagyott oszlopok közé (melyeket potenciálisan transzformálni kell), akkor gyakran nagyon sok lépésen keresztül ott is marad. Ezért lesznek olyan oszlopok, melyeket túl sokszor kell transzformálnunk.

Markowitz módszere

Markowitz pivotválasztó stratégiája szintén régóta használatos mind PFI, mind EFI invertálási technikákban [9]. Az eljárás lényege, hogy ha az aktív al mátrixban az i sor sorszámlálója R_i , a j oszlop oszlopszámlálója pedig C_j , akkor az (i, j) elemen végzett pivotálás után az aktív al mátrixban keletkező új nemnulla elemek számát felülről lehet becsülni az ún. Markowitz-számlálóval, melynek definíciója

$$(R_i - 1) \cdot (C_j - 1) .$$

A feltöltődés elkerülése érdekében a lehetséges pivotelemek közül azt választjuk, amely Markowitz-számlálója minimális. Azaz az s lépésben a következőképpen

választjuk ki az (i_s, j_s) pivotpozíciót:

$$(R_{i_s}, C_{j_s}) = \min\{(R_i - 1) \cdot (C_j - 1)\} \quad (i, j) \in I^s \times J^s,$$

ahol I^s az s -edik lépésnél a szabad sorok, J^s pedig az s -edik lépésnél szabad oszlopok indexeinek halmaza, R és C pedig az aktív almatrix sor- ill. oszlopszámlálói. A pivotáláskor, az új aktív almatrix kiszámítása közben az új sor- és oszlopszámlálók számítását is el kell végezni, majd a pivotválasztási elvet ismételten alkalmazni. Természetesen lehetőség van csak azon elemek közti pivotkeresésre, melyek eleget tesznek a stabilitási kritériumnak. Ezáltal egy olyan eljárást kapunk, mely numerikusan stabil, kis elemszámú eta file-t generál. Az egyetlen hátrány a viszonylagos lassúság, melyet az okoz, hogy minden pivotelemválasztásnál az összes lehetséges pivotelemre ki kell számítani a Markowitz-számlálót, és el kell végezni a stabilitási tesztet.

Suhl módszere

Suhl módszere [15] a Markowitz-elv olyan módosítása, mely a jó tulajdonságok megtartása mellett gyorsan működik. A módszer azon a feltételezésen alapul, hogy a minimális Markowitz-számláló olyan pozícióban található, melynek sor- illetve oszlopszámlálója eleve kicsi, ezért elegendő csak az ilyen sorokat és oszlopokat vizsgálni. Suhl dolgozata [15] EFI típusú algoritmus részeként írja le a pivotkereső eljárást.

További jelölések:

- m : a bázis rendje;
- $C_k^s = \{j \mid j \in J^s, \text{ és } B^s \text{ } j \text{ oszlopa } k \neq 0 \text{ elemet tartalmaz}\};$
- $R_k^s = \{i \mid i \in I^s, \text{ és } B^s \text{ } i \text{ sora } k \neq 0 \text{ elemet tartalmaz}\};$
- $M(j, s)$: azon b_{ij}^s , $i \in I^s$, elemekre vonatkozó minimális Markowitz-számláló, melyek kielégítik a stabilitási kritériumot, ill. ha nincs ilyen elem, akkor $M(j, s) = m^2$;
- $M(i, s)$: hasonlóan a minimális Markowitz-számláló a b_{ij}^s , $j \in J^s$ azon elemein, melyek a stabilitási kritériumnak eleget tesznek.
- p : paraméter, értéke azt mutatja, hány sor, ill. oszlop vizsgálata után hagyjuk abba a további vizsgálatokat.

Suhl pivotkereső algoritmus a következő:

FINDPIVOT algoritmus

1. Ha $C_1^s \neq 0$, pivotáljunk ezen 1 oszlopszámlálójú oszlopban, menjünk a 10. lépésre.
2. Ha $R_1^s \neq 0$, pivotáljunk ezen 1 sorszámlálójú sorban, menjünk a 10. lépésre.
3. Legyen $MM = m^2$, $n = 0$, $k = 1$.
4. Legyen $k = k + 1$.

5. Ha $k > m$, akkor menjünk a 9. lépésre.
6. Ha $C_k^s \neq \emptyset$, akkor minden $j \in C_k^s$ -ra
 - a) ha $M(j, s) < MM$, akkor
 - (i) legyen $MM = M(j, s)$,
 - (ii) legyen i a megfelelő pivotsor indexe.
 - (iii) Őrizzük meg (i, j) -t, mint az eddigi legjobb pivotpozíciót.
 - (iv) Ha $MM \leq (k-1)^2$, akkor menjünk a 10. lépésre.
 - b) Legyen $n = n + 1$.
 - c) Ha $n \geq p$ és $MM < m^2$, akkor menjünk a 10. lépésre.
7. Ha $R_k^s \neq \emptyset$, akkor minden $i \in R_k^s$ -ra
 - a) ha $M(i, s) < MM$, akkor
 - (i) legyen $MM = M(i, s)$,
 - (ii) legyen j a megfelelő pivotoszlop indexe.
 - (iii) Őrizzük meg (i, j) -t mint az eddigi legjobb pivotpozíciót.
 - (iv) Ha $MM \leq k(k-1)$, akkor menjünk a 10. lépésre.
 - b) Legyen $n = n + 1$.
 - c) Ha $n \geq p$, és $MM < m^2$, akkor menjünk a 10. lépésre.
8. Menjünk a 4. lépésre.
9. Nincs elfogadható pivot, STOP.
10. Az (i, j) pivotelem átadása, STOP.

Az algoritmus 6.a)(iv) illetve 7.a)(iv) pontján akkor lépünk ki, ha a talált pivotelemhez tartozó Markowitz-számláló minimális, azaz biztosan nem találunk jobbat, a 6.c és 7.c ponton pedig akkor lépünk ki, ha van megfelelő pivotelem és már legalább p oszlopot és sort megvizsgáltunk. A sor- és oszlopszámlálók láncolt listával való ábrázolása csekély memóriafelhasználás árán lényegesen meggyorsítja a megfelelő sorok és oszlopok kiválasztását. Suhl dolgozatában $p = 2$ és $p = 5$ értékekkel dolgozott, a pivotkeresés sebessége mindkét esetben nagyon jónak bizonyult, és az esetek túlnyomó részében az algoritmus a minimális Markowitz-számlálójú pivotelemet találta meg, így a feltöltődés is kedvezően alakult.

Egy új módszer

Hellerman és Rarick módszerében a kedvezőtlen numerikus viselkedést vélhetően az okozza, hogy néhány oszlopot az algoritmus során túl sokszor kell transzformálni, azaz ugyanazon elemeken túl sok műveletet kell végezni, és a számítási hibák okozta pontatlanság felhalmozódása ezáltal nagyon nagymérvű. Ez adta az ötletet egy olyan pivotválasztási eljáráshoz, amely megpróbálja ezt a hatást kiküszöbölni, azaz egy pivotelem kiválasztásánál figyelembe venni, hogy a régebbi pivotlépések milyen hatással voltak a bázis oszlopaira. Emellett a Markowitz-számláló részleges használatával érhető el a feltöltődés elkerülése. A következő egy ilyen, PFI típusú módszer:

Bevezetjük a következő két tömböt:

CTRAN, mely minden oszlopról tárolja, hányszor kellett az invertálás aktuális lépéséig transzformálni, és

RTRAN, mely minden sorhoz hozzárendeli azon oszlopok CTRAN értékeinek összegét, mely oszlopok tartalmazznak az adott sorhoz tartozó nemnulla elemet.

A CTRAN tömböt a következő módon számoljuk:

Az invertálás kezdetén CTRAN értéke minden oszlopra 0.

Az s lépésnél egy (i, j) helyen pivotálásnál minden $k \in J^s$ -re:

- a) ha $b_{ik}^s = 0$ akkor $CTRAN_{új}(k) = CTRAN_{rég}(k)$,
- b) ha $b_{ik}^s \neq 0$ akkor $CTRAN_{új}(k) = CTRAN_{rég}(k) + 1$.

Az RTRAN tömböt pedig:

Az invertálás kezdetén RTRAN értéke minden sorra 0, egy tetszőleges s lépésnél pedig minden $l \in I^s$ -re:

$$RTRAN(l) = \sum_{\substack{j \in J^s \\ b_{lj}^s \neq 0}} CTRAN(j).$$

Mindkét változó számítása csak csekély munkát igényel, és az új aktív almatrix számításakor elvégezhető, továbbá memóriaigényük sem túl nagy (mindkettő m hosszú, 2 byte-os integer tömb).

A pivotkeresést ezek után Markowitz módszeréhez hasonlóan végezzük, azzal a különbséggel, hogy minden olyan sort kihagyunk a vizsgálatból, amelyre az RTRAN érték nagyobb az eddigi legjobb pivotelem sorának RTRAN értékénél. Ezáltal a pivotkeresés időigénye nem rosszabb Markowitz módszerének időigényénél, megfelelően rövid eta file keletkezik, sok esetben lényegesen rövidebb, mint Markowitz módszerével, és tapasztalataim szerint numerikusan stabilabb a működés. A módszer EFI invertálásban is használható, egyedül a sorok és az oszlopok szerepét kell felcserélni.

Ezt az algoritmust a későbbiekben INVERT3-nak fogom nevezni.

5. Számítógépes implementáció

A nagyméretű feladatok ritkás tulajdonságának kihasználása központi probléma az LP feladatok implementációjában. A feladatok nemritkán több ezer feltételből és változóból is állhatnak, kitöltöttségük pedig a legtöbb esetben egy százaléknál is kisebb, igen nagy méret esetén pedig egy ezrelék alatti is lehet. Ennek megfelelően speciális adatszerkezetek használata válik szükségessé. A megoldandó feladatok a már szabványos MPS input formátumban vannak megadva. Az első eljárás tehát az MPS formátumú feladat beolvasását és az úgynevezett ritkás formátumú tárolását végzi.

A dolgozatban az EFI és PFI típusú invertálások, illetve különböző pivotválasztási szabályok gyakorlati viselkedését teszteltük. PFI esetben a Hellerman-Rarick,

Markovitz és az Invert3 algoritmusokat teszteltük, míg EFI esetben a Suhle, Markovitz és Invert3 algoritmusokat. Az összehasonlításban a fő szempontok a futási idő, a generált eta file hossza és a numerikus megbízhatóság voltak.

Adatszerkezetek

A megoldás során előforduló összes mátrixot ritkás formátumban tároltam el, minden nemnulla elemet egy következő típusú rekord ír le:

```

elemtípus = RECORD
  index : INTEGER;
  érték : VALÓS;
END

```

Az érték a nemnulla elem értéke, az index pedig oszlopfolytonos tárolás esetén az elemhez tartozó sorindex, sorfolytonos tárolás esetén pedig az oszlopindex. A feladat mátrixát oszlopfolytonos alakban tároltam, ami a következőt jelenti:

Az m rendű bázishoz be kell vezetni egy OMMUT $(m + 1)$ elemű, 4 byte-os pointer tömböt, melynek elemei a memória azon kezdőcímére mutatnak, ahol az adott oszlop tárolása kezdődik (az $(m + 1)$ -edik elem pedig az utolsó oszlop által lefoglalt memória utáni első címet tartalmazza).

Az OMMUT által mutatott címtől kezdve be kell írni elemtípusú rekordokkal az oszlop elemeit.

Egy oszlopban a nemnulla elemek számát a két szomszédos OMMUT érték különbségéből számíthatjuk ki (ezért van szükség az $(m + 1)$ -edik elemre).

Ezt a következő példával illusztrálhatjuk :

$$A = \begin{bmatrix} 0 & 0 & 5 & 3 & 0 \\ 0 & 4 & 0 & 0 & 1 \\ 2 & 0 & 0 & 9 & 0 \end{bmatrix}$$

mátrix oszlopfolytonos tárolása a következő:

	1	2	3	4	5	6
OMMUT	1	2	3	4	6	7
MEMÓRIA	(3;2,0)	(2;4,0)	(1;5,0)	(1;3,0)	(3;9,0)	(2;1,0)

PFI algoritmusok adatszerkezete

A feladat mátrixából kimásolt bázismátrix tárolása is ugyanúgy történik, ahogy a feladat teljes mátrixának tárolása, azaz oszlopfolytonosan. Ehhez egy OBMUT tömböt kell felvenni az oszlopok kezdőcímének tárolásához, és továbbá egy OSZÁML

m elemű 2 byte-os integer tömböt, mely az oszlopszámlálót tartalmazza. Ez utóbbira azért van feltétlenül szükség, mert az algoritmus közben az oszlopokat elmozgatjuk és nem marad meg a „hézag nélküli” folytonosság, így az oszlop hosszát nem tudjuk a két szomszédos oszlopmutató értékének különbségéből számolni (ellentétben a feladat mátrixának esetével). Ezért az oszlopok hosszának a tárolása is szükséges. Az eta vektorok a bázismátrix azon oszlopának helyére kerülnek, amelyből elkészítjük azokat, és így az invertálás végén a bázismátrix helyén az eta vektorokat kapjuk, oszlopfolytonos tárolással, az egyes vektorok kezdőcímét az OBMUT mutató adja, hosszát pedig az OSZÁML változóban kapjuk. Ha a bázis egy oszlopa update-elés közben méretének megnövekedése miatt nem férne el a régi helyére, akkor ezt az oszlopot áthelyezzük a szabad memória elejére, a szabad memória kezdőcímét megfelelően megnöveljük, méretét pedig csökkentjük ezen oszlop végleges méretének megfelelően. Ezáltal „hézagok” alakulnak ki az oszlopok között, és ha a rendelkezésre álló szabad memória elfogy, ezeket a hézagokat meg kell szüntetni, azaz minden oszlopot a memóriában közvetlenül előtte található oszlop vége utáni első memóriacímre kell másolni, és természetesen az OBMUT értékét módosítani. Ennek meggyorsítása érdekében célszerű felvenni egy PERMUT nevű m hosszú 2 byte-os integer tömböt, mely az egyes oszlopok aktuális elhelyezkedését tárolja láncolt lista formájában. Ezt a láncolt listát csak akkor kell módosítani, ha egy oszlopot (az előzőekben említett okból kifolyólag) a szabad memória elejére helyezzünk. A PERMUT lista használata a hézagok megszüntetésének algoritmusát (GARBAGE COLLECTION) rendkívül egyszerűvé teszi. Célszerű továbbá az oszlopok update-elését a PERMUT által meghatározott sorrendben végezni. Ez azzal az előnnyel jár, hogy ha egy oszlopot elraktunk a helyéről, a memóriában közvetlenül utána elhelyezkedő oszlop transzformálásánál az új elemeket az előző (áthelyezett) oszlop helyére, vagyis a „hézagba” kezdetjük tölteni, miáltal ha az új oszlop elfér a „hézag+régi oszlop” memóriaterületen, akkor a „hézag” a transzformációkban következő oszlop elé vándorol. Ezt az elvet tovább használhatjuk, vagyis ha egy oszlop előtt „hézag” található, akkor a hézagba kezdjük tölteni az update-elt elemeket. Tapasztalataim szerint ezen ötlettel sok memóriaműveletet és GARBAGE COLLECTION-t meg lehet spórolni.

Az invertálás befejeztével érdemes a szabad memóriaterület növelése érdekében egy GARBAGE COLLECTION-t végezni és azon a változók által lefoglalt memóriaterületeket felszabadítani, melyeket csak az invertálásban használunk.

EFI algoritmusok adatszerkezete

EFI algoritmusoknál a bázismátrix sorfolytonos alakban való tárolására van szükség. Ehhez (az előzőekhez hasonlóan) egy SBMUT tömb szükséges a sorok kezdőcímeinek és egy SSZÁML tömb a sorok hosszának tárolásához. A feladat oszlopfolytonos mátrixából a bázis sorfolytonos alakú kimásolása három lépésben történik :

1. A bázismátrix SSZÁML értékeinek kiszámítása.
2. Helyfoglalás a soroknak és SBMUT értékeinek meghatározása.
3. A bázismátrix elemeinek beírása az előre lefoglalt helyre.

Az EFI algoritmusoknál szükség van továbbá az eta vektorok tárolására, ez oszlopfolytonos alakban történik az LMUT m elemű mutató tömb segítségével. Ezeket a vektorokat a szabad memória végére rakjuk, majd a szabad memória méretét annnyival csökkentjük, amennyi tárterületet az új eta vektor lefoglal, és a szabad memória végének címét is ennek megfelelően átállítjuk. Az U mátrix sorvektorainak tárolása a bázis soraiban történik, azaz az algoritmus végen a bázismátrix helyén az U mátrixot kapjuk, sorfolytonos tárolással. A vektorok kezdőcímét az SBMUT tömbben kapjuk, a vektorokban levő elemek számát pedig a SSZÁML tömbben. A későbbi gyorsabb műveletvégzések érdekében az U mátrixot ezután oszlopfolytonos alakra írjuk (hasonlóan három lépésben, mint az oszlop \rightarrow sorfolytonos váltásnál), közben elvégezve a háromszögmátrix invertálását.

A PERMUT tömb használata itt is ugyanolyan elvek szerint történhet, továbbá a GARBAGE COLLECTION használatára is ugyanúgy kerül sor, ahogy az előző fejezetben.

Toleranciák használata

Az elvégzendő műveletek nagy száma miatt a kerekítési hibák halmozódására lehet számítani. A legnagyobb veszélyt az jelenti, hogy ha valamely helyen egy nullának kellene állnia, de a pontatlanságok miatt ott egy más értéket kapunk. Ez különösen akkor kellemetlen, ha a nulla helyett kapott érték akkora, hogy pivotelemként is elfogadható, és ki is választjuk pivotelemnek. Ez esetben ugyanis az egész inverz használhatatlanná válik. A kerekítési pontatlanság a gépi számábrázolás függvénye, a szabványosan használatos valós számtípusok a Turbo Pascal rendszerben a következők:

Típus	ábrázolás (bit)	exponens	mantissza	értékes jegyek száma
single	32	8	23	7
real	48	8	39	11
double	64	11	39	15
extended	80	15	52	19

A gyakorlatban a leggyakrabban használatos a DOUBLE típus. Numerikusan az összeadás a legproblematisabb művelet, pontosabban az, hogy az összeadás nem asszociatív. Ez a következő példán demonstrálható (SINGLE típussal végezz

a műveleteket) [12]:

$$\begin{aligned}a &= -0,56324 \cdot 10^5, \\b &= +0,56323 \cdot 10^5, \\c &= +0,99216 \cdot 10^0\end{aligned}$$

értékekkel

$$\begin{aligned}(a + b) + c &= -0,78400 \cdot 10^{-2}, \\(b + c) + a &= -0,10000 \cdot 10^{-1}, \\(a + c) + b &= 0\end{aligned}$$

eredményeket kapunk. Ezt a jelenséget úgy próbáljuk meg kiküszöbölni, hogy összeadás helyett egy \oplus műveletet végzünk, amit a következőképpen definiálunk:

$$A \oplus B = \begin{cases} A + B, & \text{ha } \frac{|A + B|}{\max(|A|, |B|)} > T_+, \\ 0, & \text{ha } \frac{|A + B|}{\max(|A|, |B|)} \leq T_+, \end{cases}$$

ahol $0 < T_+ < 1$ egy általunk választott tolerancia. Ez azt jelenti, hogy egy kivonás eredményét nullának tekintjük, ha olyan számokat vonunk ki egymásból, melyek első p értékes jegye megegyezik, ahol $p = -\log_{10}(T_+)$. REAL aritmetikánál T_+ értékét 10^{-6} -nak, DOUBLE aritmetikánál pedig 10^{-11} -nek választottam, figyelembe véve a két különböző aritmetikánál az értékes jegyek számát [12].

Ettől kissé eltérő módon célszerű a skaláris szorzást elvégezni [10]. Egy $C = V \cdot W$ skaláris szorzásnál először külön a pozitív és külön a negatív tagokat adjuk össze a szokásos összeadással, majd a kapott C^+ és C^- mennyiségeket az előbbi \oplus művelettel, azaz:

$$\begin{aligned}C^+ &= \sum_{i|\text{sign}(V_i)=\text{sign}(W_i)} V_i \cdot W_i, \\C^- &= \sum_{i|\text{sign}(V_i)\neq\text{sign}(W_i)} V_i \cdot W_i, \\C &= C^+ \oplus C^-.\end{aligned}$$

A 4. fejezetben tárgyalt stabilitási kritériumhoz választott toleranciát Tpivot1 néven használtam, illetve PFI algoritmusok második fázisú pivotkereséséhez a Tpivot2-t. Értékeik:

$$\begin{aligned}\text{Tpivot1} &= 10^{-3}, \text{ és} \\ \text{Tpivot2} &= 10^{-6}.\end{aligned}$$

Az invertálás újakezdésénél az új pivotválasztási toleranciát szintén a 4. fejezetben leírt

$$\text{Tpivot1}_{\text{új}} = \min \{2 \cdot \text{Tpivot1}_{\text{rég}} , 1.0\}$$

előírással határoztam meg.

Pivotkeresés

Ebben a fejezetben az általam vizsgált 6 pivotkereső algoritmushoz szükséges többletváltozók memóriaigényét fogjuk összehasonlítani. Mivel mind a hat esetben m elemű, 2 byte-os integer tömbökre van szükség, a későbbiekben ezt nem jelölöm külön-külön (m a bázis rendje). PFI algoritmus esetében az OSZÁML oszlopszámlálókat, EFI algoritmus esetében pedig a SSZÁML sorszámlálókat mindenképpen tárolni kell. Továbbá egy MAXMUT tömbre is szükség van mind a 6 esetben az oszlopok (EFI esetben sorok) maximális elemének indexeléséhez, ami a stabilitási kritérium gyors ellenőrzéséhez szükséges. Ezért csak az ezeken kívüli változókat soroljuk fel.

1. HELLERMAN-RARICK (PFI)

TKN : a $t_k(n)$ értékek számításához,

INS : az S -hez tartozó oszlopok nyilvántartásához,

SSZ : a nem S -beli oszlopokhoz tartozó sorszámlálók,

OSZ : az S -beli oszlopokhoz tartozó sorszámlálók, azaz összesen négy tömb.

2. MARKOWITZ (PFI)

SSZ : az aktív almatrix sorszámlálói, összesen egy tömb.

3. INVERT3 (PFI)

SSZ : az aktív almatrix sorszámlálói,

RTRAN, CTRAN : a 4 fejezetben leírtak szerint, összesen három tömb.

4. SUHL (EFI)

OSZ : az aktív almatrix oszlopszámlálói,

SLF : sorszámlálók előre-láncolt listája,

SLB : sorszámlálók hátra-láncolt listája,

SLF : sorszámlálók láncolásának kezdőpont-indexei,

OLF : oszlopszámlálók előre-láncolt listája,

OLB : oszlopszámlálók hátra-láncolt listája,

OLF : oszlopszámlálók láncolásának kezdőpont-indexei, összesen hét tömb. A sor- és oszlopszámlálók mindkét irányú láncolására azért van szükség, hogy a transzformációk végzése közben az előre-láncot gyorsan update-elhessük.

5. MARKOWITZ (EFI)

OSZ : az aktív almatrix oszlopszámlálói, összesen egy tömb.

6. INVERT3 (EFI)

OSZ : az aktív almatrix oszlopszámlálói,
 RTRAN, CTRAN : a 4 fejezetben leírtak szerint, összesen három tömb.

Transzformációk a bázisinverzzel

A simplex algoritmus fő lépései a következők [6]:

1. A simplex szorzó számítása: $\pi = v \cdot B^{-1}$, ahol v a célfüggvény sorára mutató egységvektor. Ez a BTRAN művelet.
2. A redukált költségek számítása minden nem bázisbeli a_j oszlopra: $d_j = \pi a_j$. Ez a PRICE művelet.
3. Néhány q oszlop bázisbeli képének számítása, melyekre $d_q < 0$: $\alpha_q = B^{-1}a_q$. Ez az FTRAN művelet.
4. Legyen $\beta = B^{-1}b$, és határozzuk meg a p pivotsort:

$$\frac{\beta_p}{\alpha_p} = \min_{\alpha_i > 0} \left(\frac{\beta_i}{\alpha_i} \right).$$

Ez a CHUZR művelet.

5. A bázisinverz és β update-elése, ez a WRETA művelet.

Az inverzzel FTRAN és BTRAN műveleteket kell végezni. Megjegyezzük, hogy a BTRAN sokkal pontatlanabb, mert skaláris szorzásokat kell számolni. Egy elemi FTRAN művelet a következő:

$$\begin{bmatrix} 1 & & & \eta_1 & & \\ & \ddots & & \vdots & & \\ & & & \eta_s & & \\ & & & \vdots & \ddots & \\ & & & \eta_m & & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_s \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} b_1 + b_s \eta_1 \\ \vdots \\ b_s \eta_s \\ \vdots \\ b_m + b_s \eta_m \end{bmatrix},$$

egy elemi BTRAN pedig :

$$[b_1, \dots, b_m] \begin{bmatrix} 1 & & & \eta_1 & & \\ & \ddots & & \vdots & & \\ & & & \eta_s & & \\ & & & \vdots & \ddots & \\ & & & \eta_m & & 1 \end{bmatrix} = \left[b_1, \dots, \sum_{i=1}^m b_i \eta_i, \dots, b_m \right].$$

Az algoritmusaink által előállított inverzek pontosságát FTRAN és BTRAN műveletekkel ellenőrizhetjük, mégpedig a következőképpen:

1. FTRAN CHECK : számítsuk ki minden bázisbeli b oszlopra a $B^{-1}b$ szorzatot. Végeredményül egy egységvektort kellene kapnunk, és így a kapott eredmény

és az egységvektor különbsége egy perturbációs vektor lesz, ezen perturbációs oszlopvektorokat összerakva kapjuk az E_{FTRAN} perturbációs mátrixot, azaz $B^{-1}B = I + E_{FTRAN}$.

2. BTRAN CHECK : kiszámítjuk a szimplex szorzót, majd a bázisbeli változók redukált költségeit. Itt az E_{BTRAN} perturbációs vektort a redukált költségek-ből összerakott vektor adja, ugyanis a bázisbeli változók redukált költségének pontos értéke nulla.

Az E_{FTRAN} mátrix, és az E_{BTRAN} vektor normáinak összegét a függelék 4. táblázata tartalmazza. (A táblázatokban mindenhol végtelen norma szerepel, és a kondíciós szám is a végtelen normához tartozik.)

Számítástechnikai tapasztalatok

A programokat Pascal 5.5 programozási nyelven készítettem és a Turbo Pascal 5.5 compiler következő opcióival fordítottam:

\$A+,B-,D+,E-,F-,I+,L+,N+,O-,R-,S-,V+

\$M 16384,0,655360

Futtatáshoz egy IBM PC 486/25 Mhz gépet használtam. A teljes forrásnyelvi szöveg programonként kb. 2500 sor, az impementáció elsődleges célja a bemutatott algoritmusok több szempont szerinti kényelmes vizsgálata volt. Tesztfeladatként a Maros István által összeállított, 6 különböző LP feladatból vett összesen 20, numerikusan rosszul viselkedő bázist használtam. A függelék első táblázata tartalmazza a bázisok, és explicit inverzük fontosabb adatait. A legfontosabb szempontok mérőszámai a függelék 2-4 táblázataiban találhatóak, a második táblázat a különböző algoritmusok által készített eta file nemnulla elemeinek számát tartalmazza, a harmadik táblázat a futási időket, a negyedik pedig az FTRAN és BTRAN perturbációs normák összegét.

1. Az eta file elemszámának szempontjából az EFI algoritmusok sokkal hatékonyabbnak bizonyultak a PFI algoritmusoknál, általában feleakkora eta file keletkezett az EFI algoritmus eredményeként. Ennek fő oka az aktív almatrix eltérő definíciója, azaz hogy EFI esetben azon sorokban levő elemeket nem kell transzformálni az új aktív almatrix számításakor, melyekben már pivotáltunk, PFI esetben viszont kell (kivéve a már pivotált oszlopokban levő elemeket), ezáltal több műveletet kell végezni, és így a feltöltődés is nagyobb. Az EFI esetben számottevő különbség az eta file elemszámában nem volt, míg PFI esetben az új INVERT3 algoritmus sok esetben jelentősen rövidebb eta file-t generált. (Pl. a STAIR bázisainál 1000-2000 elemmel kevesebbet, mint a másik két PFI algoritmus.)

2. A futási időben is hasonló eredmények születtek: az EFI algoritmusok 30-50%-kal gyorsabbak a PFI algoritmusoknál. (Ennek fő oka az előbb említett kevesebb műveletvégzés és kisebb feltöltődés). Az EFI algoritmusok közül a leggyorsabb Suhl algoritmus volt, a másik kettő gyakorlatilag azonos időt igényelt. Alaposabb vizsgálatok azt mutatták, hogy Suhl algoritmusánál a pivotkeresésre fordított idő az összi idő kb. 8%-át tette ki, és bár a másik két algoritmus pivotkeresése 2-3-szor

is lassabb, ez a kis részarány miatt az összes futási időben nem okozott döntő különbséget.

3. Numerikusan szintén az EFI algoritmusok bizonyultak jobbnak. A PFI algoritmusok közül Hellerman-Rarick módszere feltűnően sok esetben egyszerűen csődöt mond, a perturbációk nemegyszer olyan nagymérvűek, hogy megállapítható, az inverz teljesen használhatatlan. (Pl. a STAIR feladat bázisainál.) Ennek oka, hogy a maximális CTRAN értékek legalább kétszer akkorák, mint a másik két algoritmusnál, és ez azt eredményezi, hogy azon oszlopokban nagyon felnőnek az elemek, amelyeket túl sokszor transzformálunk. Amikor aztán egy ilyen oszlopban keresünk pivotelemet, kénytelenek vagyunk túl nagy pivotelemet választani. A maximális és minimális pivotelemek hányadosa nagyon felnő (a STAIR[5]-nél pl. ez az arány 10^{17} volt), ami nyilvánvalóan numerikusan rendkívül kedvezőtlen. A 4. táblázat alapján az INVERT3 numerikusan jobban viselkedik a Markowitz módszerénél, az INVERT3 esetében sehol sem kapunk 10^{-9} -nél nagyobb perturbációt, míg a Markowitz-módszernél a legnagyobb perturbáció 10^{-4} nagyságrendű.

Az EFI algoritmusok FTRAN művelete gyakorlatilag minden esetben tökéletesen pontos volt. A táblázatokban szereplő 10^{-16} nagyságrendű értékek csupán számkonvertálási műveletek eredménye, a DOUBLE aritmetika értékes jegyeinek száma ugyanis 15, és perturbációt csak a diagonálisban, az 1 értékű elemtől való eltérés okozott. Ugyanez a helyzet REAL aritmetika használatánál, ott az értékes jegyek száma 11, és a keletkező perturbációs értékek 10^{-13} nagyságrendűek. DOUBLE aritmetika használatával a BTRAN műveletek is nagyon pontosak voltak. Csak 3-4 bázisnál keletkezett eltérés, az INVERT3 látszik a legjobbnak, a maximális perturbáció itt 10^{-12} nagyságrendű, míg a másik két eljárásnál ez 10^{-7} illetve 10^{-9} nagyságrendű. REAL aritmetikánál azonban szembeötlőbbek a különbségek. Suhl módszerével néhány esetben (Pl. GROW22[2], STAIR[4], STAIR[5], PILOT4[3], PILOT4[7]) nem kaptam kielégítő eredményt, a maximális perturbációs norma 10^{+2} volt. Hasonlóan a Markowitz algoritmus sem volt eredményes néhány esetben (a maximális perturbáció értéke 10^{-1} nagyságrendű). Az INVERT3 meglepően jó eredményeket adott, összesen két bázisnál keletkezett perturbáció, mégpedig a PILOT4[2]-nél 10^{-8} , a PILOT4[6]-nál pedig 10^{-9} nagyságrendű. Figyelemre méltó továbbá, hogy az INVERT3-nál a maximális és minimális pivotelemek hányadosa majdnem minden esetben legalább 1 nagyságrenddel jobbnak (azaz kisebbnek) bizonyult, mint a másik két algoritmusnál tapasztalható érték.

4. A rövidebb eta file folytán az EFI algoritmusok használata gyorsabbnak bizonyult az FTRAN és BTRAN műveleteknél, mint a PFI algoritmusok alkalmazása.

Tapasztalataimat összefoglalva, az EFI algoritmusok mind a négy vizsgálati szempont alapján sokkal kedvezőbbek a PFI algoritmusoknál, akár a Bartels-Golub, akár a Forrest-Tomlin bázisinverz update-elési módszert alkalmazva minden tekintetben nagyon jó invertálási eljárást kapunk. Az INVERT3 egy olyan pivotválasztási eljárásra tett próbálkozás volt, mely megpróbálja figyelembe venni pivotkereséskor a régebbi pivotálások numerikus hatását. A vizsgált bázisokra az új elgondolás biztató eredményeket adott, melyet a módszer finomításával valószínűleg javítani

lehetne.

További vizsgálatokat igényelne az a megfigyelés, hogy az egyes invertáló eljárások tulajdonságai nagy mértékben függenek a bázistól, azaz vélhetően a bázis és a feladat struktúrájától. Egy ilyen algoritmus először a bázis vagy a feladat struktúráját vizsgálná, majd a felderített struktúrához a korábbi tapasztalatok alapján legjobban „testre szabott” algoritmust választaná.

Függelék

Név	:	multi [1]	grow7 [1]	grow7 [2]	grow22 [1]	grow22 [2]	stair [1]	stair [2]	stair [3]	stair [4]	stair [5]
Rang	:	62	141	141	441	441	357	357	357	357	357
Nemnulla elemek száma a bázisban	:	354	1449	1431	4714	4803	3442	3422	3519	3588	3443
Max. abs. értékű elem a bázisban	:	3,471E+02	7,000E+00	7,000E+00	7,000E+00	7,000E+00	9,853E+00	9,853E+00	9,853E+00	9,853E+00	9,853E+00
Trianguláris rész	:	49	12	19	18	16	57	71	37	33	49
A bázis determinánsának logaritmusa	:	-1,997E+00	-1,971E+01	-1,540E+01	-4,277E+01	-3,942E+01	4,118E+01	3,960E+01	4,364E+01	4,478E+01	2,992E+01
A bázis normája	:	3,610E+02	8,000E+00	8,000E+00	8,000E+00	8,000E+00	2,590E+01	2,590E+01	2,590E+01	2,590E+01	2,590E+01
Az inverz normája	:	7,691E+02	1,002E+06	6,594E+03	8,025E+03	1,053E+08	9,667E+02	3,464E+03	1,061E+03	3,461E+03	1,922E+06
Nemnulla elemek száma az inverzben	:	483	12695	12127	109304	163444	105655	101426	114311	116019	83902
Max. abs. értékű elem az inverzben	:	7,396E+02	3,179E+05	4,338E+03	5,416E+03	2,652E+07	5,487E+01	2,255E+02	4,594E+01	1,399E+02	2,516E+05
A bázis kondíciószáma	:	2,776E+05	8,018E+06	5,276E+04	6,420E+04	8,421E+08	2,504E+04	8,972E+04	2,748E+04	8,965E+04	4,978E+07

Név	:	perold [1]	perold [2]	perold [3]	pilot4 [1]	pilot4 [2]	pilot4 [3]	pilot4 [4]	pilot4 [5]	pilot4 [6]	pilot4 [7]
Rang	:	626	626	626	411	411	411	411	411	411	411
Nemnulla elemek száma a bázisban	:	2853	2971	2981	3256	3236	3238	2651	3195	3149	3260
Max. abs. értékű elem a bázisban	:	2,361E+04	2,361E+04	2,361E+04	2,784E+04	2,784E+04	2,784E+04	2,784E+04	2,784E+04	2,784E+04	2,784E+04
Trianguláris rész	:	240	214	211	132	134	141	250	169	161	143
A bázis determinánsának logaritmusa	:	2,938E+02	3,183E+02	3,259E+02	2,990E+02	2,991E+02	3,335E+02	3,069E+02	2,409E+02	2,301E+02	2,744E+02
A bázis normája	:	5,811E+04	5,811E+04	5,811E+04	6,851E+04	6,851E+04	6,851E+04	6,851E+04	6,851E+04	6,851E+04	6,851E+04
Az inverz normája	:	9,783E+05	7,788E+04	1,071E+05	1,022E+04	2,922E+04	2,002E+05	8,670E+00	9,560E+05	1,395E+05	3,484E+05
Nemnulla elemek száma az inverzben	:	179422	190708	190218	103889	99301	93570	15131	59933	78780	75946
Max. abs. értékű elem az inverzben	:	5,662E+04	2,593E+04	1,004E+05	2,316E+03	9,601E+03	7,597E+04	1,736E+00	1,572E+05	9,939E+04	4,314E+04
A bázis kondíciószáma	:	5,685E+10	4,526E+09	6,222E+09	7,001E+08	2,002E+09	1,372E+10	5,940E+05	6,550E+10	9,560E+09	2,387E+10

1. táblázat
A kísérleti bázismátrixok jellemzői

Név	:	multi [1]	grow7 [1]	grow7 [2]	grow22 [1]	grow22 [2]	stair [1]	stair [2]	stair [3]	stair [4]	stair [5]
HELLERMAN PFI	:	363	2013	1894	7599	8451	9321	9067	10278	10812	10154
MARKOWITZ PFI	:	365	2170	2116	7970	8996	9651	8533	10431	13320	11608
INVERT3 PFI	:	365	2010	1882	7418	8204	7549	7990	8292	12427	9089
SUHLE EFI	:	354	2011	1868	6676	7013	4733	4564	5103	5502	4952
MARKOWITZ EFI	:	354	1967	1912	6759	7137	4820	4676	5193	5611	5063
INVERT3 EFI	:	354	1950	1931	6593	6746	4822	4599	5230	5926	5081

Név	:	perold [1]	perold [2]	perold [3]	pilot4 [1]	pilot4 [2]	pilot4 [3]	pilot4 [4]	pilot4 [5]	pilot4 [6]	pilot4 [7]
HELLERMAN PFI	:	8217	12373	6947	7295	7137	7228	5784	7310	7330	8162
MARKOWITZ PFI	:	8686	10383	9664	9271	8930	8868	5109	9130	8821	10290
INVERT3 PFI	:	8250	13064	9227	8932	8618	8370	5140	9065	8875	9133
SUHLE EFI	:	3687	4036	4060	4303	3850	3926	2837	4004	4079	4181
MARKOWITZ EFI	:	3746	4034	4064	4241	3845	3902	2818	4035	4081	4232
INVERT3 EFI	:	3915	4257	4054	4203	4242	4114	2845	4066	3987	4186

2. táblázat

Az éta file-ban keletkező $\neq 0$ elemek számának összehasonlító táblázata (EFI-nél L+U összesen)

Név	:	multi [1]	grow7 [1]	grow7 [2]	grow22 [1]	grow22 [2]	stair [1]	stair [2]	stair [3]	stair [4]	stair [5]
HELLERMAN PFI	:	0,0545 Sec	1,3091 Sec	1,2000 Sec	11,9273 Sec	12,8545 Sec	10,7636 Sec	10,2000 Sec	12,0182 Sec	13,0909 Sec	11,5455 Sec
MARKOWITZ PFI	:	0,0545 Sec	1,3636 Sec	1,2545 Sec	12,7091 Sec	13,1818 Sec	8,1273 Sec	7,3636 Sec	9,0909 Sec	10,0545 Sec	8,8364 Sec
INVERT3 PFI	:	0,0545 Sec	1,2909 Sec	1,1818 Sec	12,0727 Sec	12,5818 Sec	7,7818 Sec	7,4545 Sec	8,6000 Sec	9,8909 Sec	8,5455 Sec
SUHLE EFI	:	0,0545 Sec	1,0000 Sec	0,9273 Sec	8,9273 Sec	9,1818 Sec	5,0909 Sec	4,8545 Sec	5,6181 Sec	5,8182 Sec	5,2545 Sec
MARKOWITZ EFI	:	0,0727 Sec	1,1636 Sec	1,0727 Sec	10,8364 Sec	11,1091 Sec	5,9636 Sec	5,5818 Sec	6,8182 Sec	7,2181 Sec	6,1636 Sec
INVERT3 EFI	:	0,0545 Sec	1,1091 Sec	1,0364 Sec	10,9455 Sec	11,0364 Sec	6,0364 Sec	5,6545 Sec	6,7636 Sec	7,4000 Sec	6,2545 Sec

Név	:	perold [1]	perold [2]	perold [3]	pilot4 [1]	pilot4 [2]	pilot4 [3]	pilot4 [4]	pilot4 [5]	pilot4 [6]	pilot4 [7]
HELLERMAN PFI	:	12,0545 Sec	14,3636 Sec	12,1818 Sec	7,2545 Sec	7,2364 Sec	6,9818 Sec	4,5273 Sec	6,5273 Sec	6,4364 Sec	7,2909 Sec
MARKOWITZ PFI	:	11,3091 Sec	12,5455 Sec	12,3636 Sec	7,6545 Sec	7,4000 Sec	6,9091 Sec	3,9091 Sec	6,5818 Sec	6,7818 Sec	7,6545 Sec
INVERT3 PFI	:	11,6909 Sec	13,9091 Sec	12,8182 Sec	7,6364 Sec	7,5091 Sec	7,0545 Sec	4,000 Sec	6,8727 Sec	6,7818 Sec	7,4909 Sec
SUHLE EFI	:	9,1273 Sec	9,6364 Sec	9,3273 Sec	5,3091 Sec	5,2909 Sec	4,9818 Sec	2,7818 Sec	4,2182 Sec	4,3636 Sec	4,9636 Sec
MARKOWITZ EFI	:	10,1273 Sec	10,8909 Sec	10,9091 Sec	6,2909 Sec	6,0909 Sec	5,8364 Sec	3,4727 Sec	5,1636 Sec	5,3091 Sec	5,9273 Sec
INVERT3 EFI	:	10,2727 Sec	11,1273 Sec	11,1455 Sec	6,3818 Sec	6,2909 Sec	5,9455 Sec	3,6545 Sec	5,2364 Sec	5,4182 Sec	6,0182 Sec

3. táblázat
Futási idők összehasonlítása

Név	:	multi [1]	grow7 [1]	grow7 [2]	grow22 [1]	grow22 [2]	stair [1]	stair [2]	stair [3]	stair [4]	stair [5]
HELLERMAN PFI	:	6,366E-12	1,668E-08	6,366E-12	1,809E-04	1,507E+02	5,331E+03	7,570E+02	7,220E+03	1,031E+06	3,789E+02
MARKOWITZ PFI	:	3,331E-16	2,665E-15	5,551E-16	1,819E-12	5,440E-10	2,595E-13	8,015E-13	2,132E-14	1,999E-12	2,218E-13
INVERT3 PFI	:	6,661E-16	1,341E-16	8,882E-16	4,091E-11	3,639E-13	5,738E-12	4,485E-14	2,081E-14	1,916E-12	2,731E-13
SUHLE EFI	:	1,110E-16	1,110E-16	1,110E-16	6,147E-07	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16
MARKOWITZ EFI	:	1,110E-16	1,110E-16	1,110E-16	1,399E-09	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16
INVERT3 EFI	:	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16	1,110E-16
EFI ARITMETIKA : REAL											
SUHLE EFI	:	9,095E-13	9,095E-13	9,095E-13	6,433E+02	9,095E-13	9,095E-13	9,095E-13	9,095E-13	5,212E-02	4,930E-05
MARKOWITZ EFI	:	9,095E-13	1,374E-02	9,095E-13	3,560E-07	9,095E-13	9,095E-13	9,095E-13	9,095E-13	1,137E-01	9,095E-13
INVERT3 EFI	:	9,095E-13	9,095E-13	9,095E-13	9,095E-13	9,095E-13	9,095E-13	9,095E-13	9,095E-13	9,095E-13	9,095E-13

Név	:	perold [1]	perold [2]	perold [3]	pilot4 [1]	pilot4 [2]	pilot4 [3]	pilot4 [4]	pilot4 [5]	pilot4 [6]	pilot4 [7]
HELLERMAN PFI	:	1,954E-03	7,182E+02	1,773E-03	2,429E+00	3,061E-04	4,120E-01	1,403E-06	4,145E-02	8,294E-04	3,484E-06
MARKOWITZ PFI	:	6,251E-15	1,067E-14	4,563E-13	4,666E-06	1,046E-04	4,578E-13	7,105E-15	1,384E-11	1,617E-11	1,514E-08
INVERT3 PFI	:	4,562E-15	2,887E-09	6,894E-12	4,190E-09	1,055E-09	1,517E-13	7,105E-15	2,726E-12	1,495E-11	8,499E-10
SUHLE EFI	:	1,110E-16	7,706E-10	1,110E-16	1,110E-16	3,019E-14	8,204E-10	1,110E-16	2,105E-10	4,705E-13	1,110E-16
MARKOWITZ EFI	:	1,110E-16	1,110E-16	1,110E-16	1,110E-16	5,371E-16	1,484E-12	1,110E-16	1,170E-16	7,087E-13	1,110E-16
INVERT3 EFI	:	1,110E-16	1,110E-16	1,110E-16	1,110E-16	5,128E-15	3,931E-16	1,110E-16	1,119E-16	4,439E-12	1,110E-16
EFI ARITMETIKA : REAL											
SUHLE EFI	:	9,095E-13	9,095E-13	9,095E-13	9,095E-13	5,590E-08	7,216E-04	9,095E-13	9,095E-13	3,895E-05	3,607E-04
MARKOWITZ EFI	:	1,858E-08	9,095E-13	9,095E-13	9,095E-13	2,838E-05	9,095E-13	9,095E-13	9,095E-13	2,947E-09	3,608E-04
INVERT3 EFI	:	9,095E-13	9,095E-13	9,095E-13	9,095E-13	5,583E-08	9,095E-13	9,095E-13	9,095E-13	3,459E-09	9,095E-13

4. táblázat

Numerikus eredmények: perturbációs normák összege

IRODALOM

- [1] BARTELS, R.H. and GOLUB, G.H., „The simplex method of linear programming using LU decomposition”, *Communications ACM* 12 (1969), 266–268, 275–278.
- [2] BARTELS, R.H., „A stabilization of the simplex method”, *Numerische Mathematik* 16 (1971), 414–434.
- [3] BRAYTON, R.K. and GUSTAVSON, F.G., *Some results on sparse matrices* (RC-2332, IBM Research Centre, Yorktown Heights, N.Y., 1969).
- [4] MC.BRIDE, R.D., „The factorization approach to large-scale linear programming”, *Mathematical programming* 10 (1976), 91–110.
- [5] MC.BRIDE, R.D., „A spike collective dynamic factorization algorithm for the simplex method”, *Management Science* 24 (1978), 1031–1042.
- [6] FORREST, J.J.H. and TOMLIN, J.A., „Updated triangular factors of the basis to maintain sparsity in the product form simplex method”, *Mathematical Programming* 2 (1972), 263–278.
- [7] GOLDFARB, D., „On the Bartels-Golub decomposition for linear programming bases”, *Mathematical Programming* 13 (1977), 272–279.
- [8] HELLERMAN E. and RARICK D., „Reinversion with the preassigned pivot procedure”, *Mathematical Programming* 1 (1971), 195–216.
- [9] MARKOWITZ, H.M., „The elimination form of inverse and its application to linear programming”, *Management Science* 3 (1957), 255–269.
- [10] MAROS, I., *MILP linear programming optimizer for personal computers under DOS* (Research Report, MTA SZTAKI, Budapest, 41/1991).
- [11] MAROS, I., „Adaptív elemek a lineáris programozásban”, *Alkalmazott Matematikai Lapok* 2 (1976), 377–393.
- [12] MAROS, I., „Adaptív elemek a lineáris programozásban II”, *Alkalmazott Matematikai Lapok* 7 (1981), 1–71.
- [13] REID, J.K., „A sparsity-exploiting variant of the Bartels-Golub decomposition for linear programming bases”, *Mathematical Programming* 24 (1982), 55–69.
- [14] SAUNDERS, M.A., „The complexity of LU updating in the simplex method, in: Anderssen, R.S. and Brent, R.P., (eds.),”
- [15] SUHL, U.H. and SUHL, M.L., „Computing sparse LU factorizations for large-scale linear programming bases”, *ORSA Journal of Computing* 2 (1990), 325–335.
- [16] WILKINSON, J.H., *Rounding errors in algebraic process* (Prentice-Hall, Englewood Cliffs, N.J., 1963).

(Beérkezett: 1992. június 4.)

MÉSZÁROS CSABA
SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZET
1111 BUDAPEST, KENDE U., 13–17.

BASIS INVERSE OF LARGE-SCALE AND
SPARSE LINEAR PROGRAMMING PROBLEMS
CS. MÉSZÁROS

This paper gives a brief survey on the methods for computing and updating the basis inverse of large-scale and sparse linear programming problems. In addition, we present a new method for pivot searching which takes the numerical effects of the previous transformations also into account. We tested the different methods on numerically instable bases arising in practical problems. Our expectation on the better numerical behaviour of the new pivot searching strategy has been justified by the computational experiments.

LINEÁRIS TÖMEGPONTRENDSZER CSILLAPÍTOTT MOZGÁSAIRÓL

WIANDT TAMÁS

Szeged

A dolgozat témája mechanikai rendszerek egyensúlyi helyzetének aszimptotikus stabilitás-vizsgálata. Az általunk tanulmányozott rendszerek mozgásegyenletei

$$A\ddot{q} + B(t)\dot{q} + Cq = 0$$

alakúak, ahol $A, C \in \mathbb{R}^{n \times n}$ konstans, $B(t)$ időtől függő $n \times n$ -es mátrix. Először összefoglaljuk a már ismert eredményeket az egyszerűbb esetekre, ezután foglalkozunk az általános esettel: elegendő feltételeket adunk arra, hogy egyenletünk 0-megoldása aszimptotikusan stabilis legyen.

Bevezetés

Egy mechanikai rendszer állapotát helyzete és sebessége határozza meg. Egy ilyen rendszer adott egyensúlyi helyzetének stabilitását úgy szokás definiálni, hogy a rendszer mozgása során az állapothatározók tetszőlegesen kis mértékben térnek csak el a nyugalmi állapottól, feltéve, hogy a mozgás az egyensúlyi helyzethez elég közelről, elég kis sebességgel indult. Az egyensúlyi helyzet aszimptotikusan stabilis, ha az állapothatározók eltérése az egyensúlyi helyzettől nullához tart. Nem szükséges hangsúlyozni, mennyire fontos a gyakorlatban egy adott rendszer egyensúlyi helyzetének stabilitását és aszimptotikus stabilitását vizsgálni. (Gondolhatunk itt gépek mozgására, épületek, hidak kilengésére stb.)

A mechanikai rendszerek mozgásegyenletei általában nem integrálhatók, ezért a mozgások stabilitási tulajdonságait a megoldások ismerete nélkül, csupán az egyenletben szereplő függvények vizsgálatával kell megállapítanunk.

Az általunk tanulmányozott rendszerek mozgásegyenletei

$$(1) \quad A\ddot{q} + B(t)\dot{q} + Cq = 0$$

alakúak, ahol A, C $n \times n$ -es konstans mátrixok, $B(t)$ időtől függő $n \times n$ -es mátrix. Itt a $-B(t)\dot{q}$ a súrlódási erő, $B(t)$ a súrlódási együtthatók mátrixa. Az alapprobléma az, hogy milyen súrlódás hatására válik az egyensúlyi helyzet aszimptotikusan stabilissá. A stacionárius esetre (mikor a $B(t) \equiv B$ konstans) már igen sok eredmény ismert. A súrlódás nélküli esetben a mozgás harmonikus rezgések lineáris kombinációja, ebben az esetben az energiamegmaradás miatt nincs aszimptotikus stabilitás. Eddig szinte kizárólag skaláris esetben tanulmányozták az ilyen rendszereket. Világossá vált, hogy skaláris esetben a súrlódási együtthatót valamilyen

integrális értelemben alulról és felülről korlátozni kell az aszimptotikus stabilitás eléréséhez. Tudniillik, ha a súrlódási együttható kicsi, akkor a rendszer hasonló a súrlódásmentes rendszerekhez és így nem létezik aszimptotikus stabilitás, másrészt ha a súrlódási együttható túl gyorsan nő, a csillapítás olyan nagy lehet, hogy a rendszer nem tud visszatérni egyensúlyi helyzetéhez.

A dolgozatban ezt a kérdést vizsgáljuk rendszerek esetén. Bebizonyítjuk, hogy, durván szólva, ha a súrlódási együtthatók $B(t)$ mátrixának legnagyobb sajátértéke legfeljebb t -vel arányosan növekszik az időben és a legkisebb sajátérték konstans felett marad, akkor létezik aszimptotikus stabilitás. (Ez egyébként Z. ARTSTEIN és E. F. INFANTE eredményének általánosítása rendszerekre.)

Előzmények

Mechanikai rendszer mozgása leírható a Lagrange-féle másodfajú mozgásegyenletekkel:

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}} \right) - \frac{\partial T}{\partial q} = Q(q, \dot{q}).$$

Itt $q = (q_1, \dots, q_n)$ az általánosított koordináták, $Q(q, \dot{q})$ az általánosított erő, $T(q, \dot{q}) = (\dot{q}, A(q)\dot{q})/2$ pedig a kinetikai energia. Ha egy mechanikai rendszer stabilis egyensúlyi helyzetétől közlelről indul kis sebességgel, akkor a mozgás során az egyensúlyi helyzettől való eltérés kicsi marad. Ez lehetővé teszi, hogy a megfelelő függvényekben elegendő csak az elsőfokú tagokkal foglalkozni, a hely- és sebességkoordináták magasabb hatványait elhanyagolhatjuk. A differenciálegyenlet-rendszerünk így lineáris lesz. Az ezzel a redukcióval kapott egyenletet szokás a „kis rezgések egyenletének” nevezni, amely a következő alakot ölti:

$$A\ddot{q} + B_1\dot{q} + C_1q = 0.$$

Vezessük be a következő jelöléseket:

$$\begin{aligned} B &= (B_1 + B_1^T)/2, & G &= (B_1 - B_1^T)/2, \\ C &= (C_1 + C_1^T)/2, & D &= (C_1 - C_1^T)/2. \end{aligned}$$

Így a következő összefüggést kapjuk:

$$(2) \quad A\ddot{q} + B\dot{q} + G\dot{q} + Cq + Dq = 0.$$

A ható erőket (pl. MERKIN [7] nyomán) osztályozhatjuk a következő módon: ez az egyenletrendszer egy mechanikai rendszer mozgását írja le a $T = (\dot{q}, A\dot{q})/2$ kinetikai energiával, $-Cq$ potenciális (vagy konzervatív) erő és $P = (q, Cq)/2$ a potenciális energia, $-B\dot{q}$ disszipatív erő, ha az $R = (\dot{q}, B\dot{q})/2$ Rayleigh-féle függvény pozitív szemidefinit, $-G\dot{q}$ giroszkopikus és $-Dq$ nemkonzervatív erő. Az R függvényt a

súrlódás által elvesztett energia mérésére használhatjuk. (Ld. [3].) A konstans együtthatós rendszerek vizsgálata már a múlt században kezdődött és nagyon sok eredmény ismeretes. Fontos kérdés (2)-vel kapcsolatban, hogy milyen disszipatív erők hatására lesz aszimptotikusan stabilis a $q = \dot{q} = 0$ stabilis egyensúlyi helyzet. Ha a súrlódás elég sok energiát „elvesz” a rendszerből, akkor ez teljesülni fog. Azt mondjuk, hogy a disszipáció teljes, ha $(\dot{q}, B\dot{q}) \geq \beta |\dot{q}|^2$, ahol β pozitív konstans. SALVADORI [3] tétele szerint, ha a potenciális energiának a $q = 0$ -ban minimuma van, a $q = 0$ egyensúlyi helyzet izolált és a disszipáció teljes, akkor a $q = \dot{q} = 0$ egyensúlyi helyzet aszimptotikusan stabilis. Lényeges, hogy a tétel csak a stacionárius esetről beszél. Ennek oka, hogy a Salvadori-tétel a BARBASIN-KRASZOVSKIJ tételt használja, amely szintén csak autonóm rendszerekre vonatkozik. Kérdés, hogy mi történik akkor, ha a súrlódási és a giroszkopikus erők az időtől is explicit módon függenek. (Pl. szakaszos fékezés.) Skaláris esetben az egyenlet

$$(3) \quad \ddot{x} + b(t)\dot{x} + c^2x = 0$$

alakú ($x \in \mathbb{R}$, $b(t) \geq 0$). Ha $b(t)$ nemnegatív, akkor az $x = \dot{x} = 0$ stabilis. (A $V = \dot{x}^2/2 + c^2x^2/2$ Ljapunov-függvény ezen egyenletre vonatkozó deriváltja $\dot{V} = -b(t)\dot{x}^2 \leq 0$.) Ha a súrlódási együttható alulról és felülről pozitív korláttal bír, azaz $0 < \underline{b} \leq b(t) \leq \bar{b}$, ahol \underline{b} és \bar{b} konstansok, akkor az egyensúlyi helyzet aszimptotikusan stabil (ld. [6]). Kérdés, hogy a korlátosság elhagyható-e valamelyik oldalról. A felső korlát nem hagyható el, mert J. HALE példája szerint az $\ddot{x} + (2 + e^t)\dot{x} + x = 0$ egyenlet megoldása az $x(t) = a(1 + e^{-t})$, amely nem tart nullához, ha $t \rightarrow \infty$. $b(t)$ növekedését tehát valamilyen módon korlátozni kell.

TÉTEL (Z. ARTSTEIN-E. F. INFANTE [1]). Ha $b(t) \geq \underline{b} > 0$ és

$$\int_0^t b(t)dt \leq Kt^2$$

a t elég nagy értékeire, akkor (3) 0-megoldása aszimptotikusan stabil.

A $b(t) \geq \underline{b} > 0$ feltétel azt biztosítja, hogy $(x(t), \dot{x}(t)) \rightarrow (c, 0)$ ($t \rightarrow \infty$) teljesül. Felvetődik a kérdés, hogy az alsó korlátot milyen megszorításokkal hagyhatjuk el.

TÉTEL (R. J. BALLIEU-K. PEIFFER [2]). Ha $b(t)$ monoton csökkenő és

$$\int_0^\infty b(t)dt = \infty,$$

akkor (3) 0-megoldása aszimptotikusan stabilis.

Megvizsgáltuk tehát a konstans együtthatós n -változós és az időtől függő súrlódással rendelkező skaláris esetet. (Az utóbbi esettel foglalkoznak még pl. [5] és [8].) Természetes módon adódik a kérdés, hogy mit mondhatunk az időben változó disszipatív erő hatása alatt álló n -változós rendszerekről.

Eredmények

Problémánk tehát a következő. Tekintsük az

$$(4) \quad A\ddot{q} + B(t)\dot{q} + Cq = 0$$

egyenletrendszert, ahol A, C szimmetrikus, pozitív definit $n \times n$ -es konstans mátrixok; $B : [0, \infty) \rightarrow \mathbb{R}^{n \times n}$ szakaszonként folytonos, szimmetrikus, pozitív szemidefinit függvény. Milyen feltételeket kell tenni $B(t)$ -re, hogy az egyenlet 0-megoldása aszimptotikusan stabilis legyen? Az egyváltozós probléma legegyszerűbb tétele úgy szól, hogy ha a súrlódás alulról és felülről korlátos, akkor az egyensúlyi helyzet aszimptotikusan stabil. Ennek analogonjaként kimondhatjuk a következő tételt:

1. TÉTEL. Legyen $B(t)$ legkisebb sajátértéke $\lambda(t)$, legnagyobb sajátértéke $\Lambda(t)$. Ha $0 < \lambda_0 \leq \lambda(t) \leq \Lambda(t) \leq \Lambda_0$, akkor a $q = \dot{q} = 0$ egyensúlyi helyzet aszimptotikusan stabil.

Bizonyítás. Tekintsük a (4) egyenletrendszerrel azonos (5) rendszert:

$$(5) \quad \dot{q} = p, \quad A\dot{p} = -Cq - B(t)p.$$

Legyen $(q(t), p(t))$ megoldása (5)-nek. Vegyük a $V(q, p) = (p, Ap)/2 + (q, Cq)/2$ Ljapunov-függvényt. Ez pozitív definit. Ekkor $v(t) = V(q(t), p(t))$ -re igaz, hogy $dv/dt = -p^T(t)B(t)p(t)$, azaz $v(t)$ nemnegatív monoton csökkenő függvény. Kaptuk, hogy a $p = q = 0$ stabilis. Másrészt

$$dv/dt \leq -\lambda(t) |p|^2 \leq -\lambda_0 |p|^2.$$

Ebből [4, 6.2.2. következmény] adódik, hogy $(q(t), p(t)) \rightarrow (u, 0)$, ahol u konstans vektor.

Integráljuk (5) második egyenletét az $[m, m+1]$ -en (m legyen pozitív egész). Átrendezés után kapjuk, hogy

$$\left| \int_m^{m+1} Cq(t) dt \right| \leq |A(p(m+1) - p(m))| + \int_m^{m+1} |B(t)| |p(t)| dt.$$

Mivel $\Lambda(t) \leq \Lambda_0$, ezért

$$\left| \int_m^{m+1} Cq(t) dt \right| \leq |A(p(m+1) - p(m))| + \Lambda_0 \int_m^{m+1} |p(t)| dt.$$

De $m \rightarrow \infty$ esetén $p(m), p(m+1) \rightarrow 0$; $\int_m^{m+1} q(t) dt \rightarrow u$ és $\int_m^{m+1} |p(t)| dt \rightarrow 0$, ha $m \rightarrow \infty$. Ezért $Cu = 0$. De C nemelfajuló, tehát $u = 0$. \square

Vegyük észre, hogy a bizonyítás egyik fő lépése az, hogy $|B(t)|$ korlátos, tehát $B(t)$ nem tudja túlságosan megnövelni a nullához tartó p -t. Tehát ha a fékező erő alulról és felülről korlátos bizonyos értelemben, a 0-megoldás aszimptotikus stabilitását itt is kapjuk. Következő tételünk Z. ARTSTEIN és E. F. INFANTE tételének általánosítása lesz differenciálegyenlet-rendszerekre.

2. TÉTEL. Tekintsük a (4) rendszert. Legyen $B(t)$ legkisebb sajátértéke $\lambda(t)$, legnagyobb sajátértéke $\Lambda(t)$. Ha $0 < \lambda_0 \leq \lambda(t)$ és

$$\int_0^t \Lambda(s) ds \leq Kt^2$$

t elég nagy értékeire, akkor a $q = \dot{q} = 0$ egyensúlyi helyzet aszimptotikusan stabil.

Tételünk bizonyításához szükség van néhány lemmára.

1. LEMMA [1]. Legyen N_0, K rögzített, továbbá a_1, a_2, \dots pozitív számok tesszőleges sorozata úgy, hogy

$$\sum_{i=1}^k a_i \leq (k + N_0)^2 K$$

bármely k esetén. Ekkor

$$\sum_{i=1}^{\infty} \frac{1}{a_i} = \infty.$$

Bizonyítás. A Schwarz-egyenlőtlenségből adódik, hogy ha b_1, \dots, b_m pozitívak és $b_1 + \dots + b_m \leq L$, akkor

$$\sum_{i=1}^m \frac{1}{b_i} \geq m^2 / L.$$

Legyen most $b_i = a_{2^n+i}$, $i = 1, 2, \dots, 2^n$. A feltevésből jön, hogy

$$\sum_{i=1}^{2^n} b_i \leq (2^{n+1} + N_0)^2 K$$

és így

$$\sum_{i=1}^{2^n} \frac{1}{b_i} \geq \frac{2^{2n}}{(2^{n+1} + N_0)^2 K},$$

de ez az alsó korlát $1/4K$ -hoz tart, ha $n \rightarrow \infty$, és így

$$\sum_{i=1}^{\infty} \frac{1}{a_i} = \frac{1}{a_1} + \sum_{n=0}^{\infty} \sum_{i=1}^{2^n} \frac{1}{a_{2^n+i}} = \infty.$$

Ezzel a lemnát beláttuk. \square

2. LEMMA. Legyen H szimmetrikus, pozitív definit $n \times n$ -es konstans mátrix. Legnagyobb sajátértékét jelölje Λ . Ekkor $(Hy, Hy) \leq \Lambda(y, Hy)$ tetszőleges $y \in \mathbb{R}^{n \times n}$ esetén.

Bizonyítás. H szimmetrikus, pozitív definit, ezért létezik sajátvektoraiból álló ortonormált bázis: e_1, \dots, e_n , ahol $He_i = \lambda_i e_i$ ($0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = \Lambda$ és λ_i a H sajátértékei.) Az y vektor előáll a báziselemek lineáris kombinációjaként: $y = \sum_{i=1}^n \xi_i e_i$. Ekkor $Hy = \sum_{i=1}^n \lambda_i \xi_i e_i$, tehát a bizonyítandó egyenlőtlenség a következő alakot ölti:

$$\sum_{i=1}^n \lambda_i^2 \xi_i^2 \leq \lambda_n \sum_{i=1}^n \lambda_i \xi_i^2$$

és ez nyilván igaz. \square

3. LEMMA. Legyen H, M szimmetrikus, pozitív definit $n \times n$ -es konstans mátrix. Legnagyobb sajátértékeiket jelölje Λ , illetve Ω . Ekkor tetszőleges $x, y \in \mathbb{R}^n$ esetén

$$(Hy, Mx) + (Mx, Hy) = 2(Hy, Mx) \leq \Lambda(x, Mx) + \Omega(y, Hy).$$

Bizonyítás. A számtani és mértani közép közötti összefüggésből kapjuk, hogy

$$\frac{1}{2}(\Lambda(x, Mx) + \Omega(y, Hy)) \geq (\Lambda(x, Mx)\Omega(y, Hy))^{1/2}.$$

A 2. Lemma és a Schwarz-egyenlőtlenség szerint viszont

$$(\Lambda(y, Hy))^{1/2}(\Omega(x, Mx))^{1/2} \geq (Hy, Hy)^{1/2}(Mx, Mx)^{1/2} \geq (Hy, Mx).$$

Ezzel a lemmát beláttuk. \square

4. LEMMA. Legyen $B : [a, b] \rightarrow \mathbb{R}^{n \times n}$ szakaszonként folytonos, szimmetrikus, pozitív definit függvény. Jelölje $\Lambda(t)$ a $B(t)$ mátrix legnagyobb sajátértékét. Ekkor tetszőleges $y : [a, b] \rightarrow \mathbb{R}^n$ folytonos függvény esetén

$$\left| \int_a^b B(t)y(t)dt \right|^2 \leq \int_a^b \Lambda(t)dt \int_a^b (y(t), B(t)y(t))dt.$$

Bizonyítás. Tekintsük az $[a, b]$ egy ekvidisztáns beosztását: $a = t_0 \leq t_1 \leq \dots \leq t_m = b$, ahol $t_i - t_{i-1} = \Delta t$ ($i = 1, \dots, m$) és írjuk fel az integrálközelítő összegeket az egyenlőtlenség mindkét oldalára. Elegendő bebizonyítani, hogy

$$\left(\sum_{i=1}^m B(t_i)y(t_i)\Delta t, \sum_{i=1}^m B(t_i)y(t_i)\Delta t \right) \leq \sum_{i=1}^m \Lambda(t_i)\Delta t \sum_{i=1}^m (y(t_i), B(t_i)y(t_i))\Delta t.$$

Mindkét oldalt Δt^2 -el egyszerűsítve a bizonyítandó állítás a következő alakot ölti a belső szorzat szétbontása után:

$$\sum_{i=1}^m \sum_{j=1}^m (B(t_i)y(t_i), B(t_j)y(t_j)) \leq \sum_{i=1}^m \sum_{j=1}^m \Lambda(t_i)(y(t_j), B(t_j)y(t_j)).$$

Az összegeket is két részre bonthatjuk:

$$\begin{aligned} \sum_{i=1}^m (B(t_i)y(t_i), B(t_i)y(t_i)) + \sum_{1 \leq i < j \leq m} 2(B(t_i)y(t_i), B(t_j)y(t_j)) &\leq \\ &\leq \sum_{i=1}^m \Lambda(t_i)(y(t_i), B(t_i)y(t_i)) + \\ &+ \sum_{1 \leq i < j \leq m} (\Lambda(t_i)(y(t_j), B(t_j)y(t_j)) + \Lambda(t_j)(y(t_i), B(t_i)y(t_i))). \end{aligned}$$

Ez pedig a 2. és 3. Lemma szerint igaz. \square

A 2. Tétel bizonyítása. Az egyenletet itt is átírjuk az (5) alakba. Mivel a legkisebb sajátérték pozitív korlát felett marad, az 1. Tétel bizonyításának első része szerint itt is igaz, hogy $(q(t), p(t)) \rightarrow (u, 0)$, ha $t \rightarrow \infty$.

Bebizonyítjuk, hogy létezik monoton növvő $n_1, n_2, \dots, n_i, \dots$ egészekből álló sorozat úgy, hogy $n_i \rightarrow \infty$ és

$$(6) \quad \int_{n_i}^{n_{i+1}} B(t)p(t)dt \rightarrow 0, \quad \text{ha } i \rightarrow \infty.$$

Tegyük fel az ellenkezőjét, azaz létezik $\delta > 0$ és n_0 úgy, hogy

$$0 < \delta \leq \left| \int_n^{n+1} B(t)p(t)dt \right|^2, \quad \text{ha } n \geq n_0.$$

A 4. Lemma szerint ekkor

$$\delta \leq \int_n^{n+1} \Lambda(t)dt \int_n^{n+1} (p(t), B(t)p(t))dt.$$

Legyen $a_n = \int_n^{n+1} \Lambda(t)dt$. Ekkor a $dv(t)/dt = -(p(t), B(t)p(t))$ azonosságot is figyelembe véve kapjuk, hogy

$$\delta \frac{1}{a_n} \leq v(n) - v(n+1)$$

ebből pedig

$$\delta \sum_{i=n_0}^{\infty} \frac{1}{a_i} \leq v(n_0),$$

vagyis $\sum_{i=n_0}^{\infty} 1/a_i < \infty$. Másrészt a tétel feltétele szerint $\sum_{i=n_0}^n a_i \leq K n^2$, amiből az 1. Lemma szerint $\sum_{i=n_0}^{\infty} 1/a_i = \infty$ következik, ami ellentmondás. Tehát (6) igaz.

Integráljuk most az (5) rendszer második egyenletét az $[n_i, n_i+1]$ intervallumon:

$$A(p(n_i+1) - p(n_i)) = -C \int_{n_i}^{n_i+1} q(t) dt - \int_{n_i}^{n_i+1} B(t)p(t) dt.$$

Felhasználva a $p(t) \rightarrow 0$, $q(t) \rightarrow u$ ($t \rightarrow \infty$) és a (6) konvergenciát, a $Cu = 0$ egyenlőséget kapjuk. C nemelfajuló, ezért $u = 0$. \square

Tehát megállapíthatjuk, hogy ha a sűrűdési együtthatók $B(t)$ mátrixának sajátértékei pozitív alsó korlát felett maradnak és nem nőnek túlságosan gyorsan, akkor a $q = p = 0$ egyensúlyi helyzet aszimptotikusan stabil.

Megoldásra vár még az a probléma, hogy mit mondhatunk abban az esetben, ha a sűrűdési együtthatók mátrixának legkisebb sajátértéke 0-hoz tart. Bizonyos feltételekkel itt is teljesül az aszimptotikus stabilitás (ld. HATVANI L. [4], 54. oldal 3.3.7. következmény), de — ez a sejtésem — ezek a feltételek a jelen dolgozat módszerét használva gyengíthetők.

IRODALOM

- [1] ARTSTEIN, Z., INFANTE, E. F., „On the asymptotic stability of oscillators with unbounded damping”, *Quart. Appl. Math.* **34** (1976), 195–199.
- [2] BALLIEU, R. J., PEIFFER, K., „Attractivity of the origin for the equation $\ddot{x} + f(t, x, \dot{x})|\dot{x}|^\alpha + g(x) = 0$ ”, *J. Math. Anal. Appl.* **65** (1978), 321–332.
- [3] HATVANI, L., „Nem-autonóm differenciálegyenlet-rendszerek megoldásának stabilitása és aszimptotikus stabilitása”, *Alkalmazott Matematikai Lapok* **5** (1979), 1–48.
- [4] HATVANI, L., „Közönséges differenciálegyenletek megoldásának stabilitásáról mechanikai alkalmazásokkal”, *Alkalmazott Matematikai Lapok* **15** (1990–91), 1–90.
- [5] HATVANI, L., TOTIK, V., „Asymptotic stability of the equilibrium of the damped oscillator”, *Differential and Integral Equations*, (megjelenés alatt).
- [6] LEVIN, J. J., NOHEL, J. A., „Global asymptotic stability of nonlinear systems of differential equations and applications to reactor dynamics”, *Arch. Rat. Mech. Anal.* **5** (1960), 194–211.
- [7] MERKIN, D. P., *Giroszkopikus rendszerek* (Nauka, Moszkva, 1974).

- [8] SMITH, R. A., „Asymptotic stability of $\ddot{x} + a(t)\dot{x} + x = 0$ ”, *Quart. J. Math. Oxford Ser.(2)* **12** (1961), 123–126.

(Beérkezett: 1992. november 6.)

WIANDT TAMÁS
BOLYAI INTÉZET
H-6720 SZEGED, ARADI VÉRTANÚK TERE 1

ON DAMPED OSCILLATION OF LINEAR MECHANICAL SYSTEMS

T. WIANDT

Sufficient conditions are given for the asymptotic stability of the equilibrium $q = \dot{q} = 0$ of the system described by the equation

$$A\ddot{q} + B(t)\dot{q} + Cq = 0,$$

where $q \in \mathbb{R}^n$, $A, C \in \mathbb{R}^{n \times n}$, and the function $B : [0, \infty) \rightarrow \mathbb{R}^{n \times n}$ is piecewise continuous. Among others, it is proved that if the smallest eigenvalue of $B(t)$ remains above a positive constant, and the largest eigenvalue $\Lambda(t)$ of $B(t)$ satisfies an inequality

$$\int_0^t \Lambda(s) ds \leq Kt^2 \quad (K = \text{const.})$$

for large values of t , then $q = \dot{q} = 0$ is asymptotically stable.

1
2
3
4
5
6

7

8

9

10
11
12
13

14

15

16

17

KOCKA ALAPTARTOMÁNYÚ EUKLIDESZI KRISTÁLYCSOPORTOK

PROK ISTVÁN

Budapest

Az [5]-ben és [6]-ban leírt algoritmus megfelelő módosításával meghatároztuk a tér kocka alaptartományú kristálycsoportjait. Ehhez egy kockát tekintünk, amelyet valamilyen jelöléssel teljesen aszimmetrikussá tettünk. Ezután az algoritmus segítségével megkeressük a tér összes lehetséges, kombinatorikusan különböző periódikus kocka kikövezését. Az így kapott (és [8]-ban felsorolt) 298 kikövezés szimmetriái egy-egy kristálycsoportot határoznak meg, pontosan azokat, amelyeknek alaptartománya lehet a kocka. Geometriai megfontolásokkal 130 ilyen kristálycsoport adódott, amelyek mindegyike a [2]-ben felsorolt 219 tércsoport egyikével izomorf. A 130 csoport felsorolása az 1. táblázatban található. A 2. táblázatban részletesen megadjuk azt a 45 kikövezést, amelyben a kocka hat lapját 3 generátor párosítja. Ezek a kikövezések 31 tércsoportot határoznak meg.

Bevezetés

Egy adott poliéderrel kapcsolatban felvethető az a kérdés, hogy kikövezhető-e vele a tér szabályos módon, pontosabban van-e a vizsgált térnek olyan diszkrét csoportja, amelynek az adott poliéder az alaptartománya. Ha ilyen csoport létezik, akkor feladatul tűzhetjük ki az ilyen csoportok megkeresését, továbbá annak meghatározását, hogy egy megfelelő csoportnak hogyan lehet alaptartománya a poliéder.

A továbbiakban elsőként ezt a kérdést fogjuk vizsgálni. Először figyelmen kívül hagyjuk a poliéder metrikus tulajdonságait, és csak az illeszkedési struktúráját adjuk meg zászlóinak (az egymáshoz illeszkedő csúcs él, lap rendezett hármaskoknak) felsorolásával. Megfogalmazunk egy algoritmust, amely egy ilyen kombinatorikusan adott poliéderhez megkeresi azokat a tércsoportokat, amelyeknek a poliéder alaptartománya. Az eredményként adódó csoportot generátoraival és definiáló relációival meghatározva kapjuk. A relációk kitevőjét szabadon választhatjuk meg. Ez a poliéder lapszögeivel függ majd össze. A kapott csoportok általában egy-egy kombinatorikusan leírható tér diszkrét csoportjai lesznek. Ezért további vizsgálatokra van szükség annak eldöntéséhez, hogy egy adott csoport a kitevők valamely megválasztásánál lehet-e valamely állandó görbületű tér diszkrét csoportja, és ha lehet, akkor a poliéder metrikusan hogyan realizálódik [5, 6].

Ezután az algoritmuson egy egyszerű módosítást hajtunk végre. Az illeszkedési struktúrán kívül megadjuk a poliéder minden éléhez az él mentén illeszkedő lapok hajlásszögét is. Ekkor a végrehajtás során (ha lehet) meg tudjuk határozni a szabad kitevőket úgy, hogy minden élet átfedés nélkül teljesen körülvegyék a poliéder

példányok. Ez az algoritmus közvetlenül csak olyan szigorúan összefüggő, metrikusan adott poliéderekre működik, amelyeknek minden lapja szabályos sokszög, és az azonos oldalszámú lapok egybevágók. Ezen algoritmus segítségével közvetlenül meg lehet határozni a poliédert tartalmazó tér megfelelő diszkrét csoportjait. A poliéder teljes metrikus leírását megadva és az algoritmust megfelelően módosítva, általános poliéderek vizsgálata is megoldható.

A kidolgozott algoritmus számítógépi programját is elkészítettük. Segítségével megkerestük az euklideszi tér kocka alaptartományú kristálycsoportjait. A vizsgálatból az az eredmény adódott, hogy a tér kockával történő szabályos kikövezése 298 kombinatorikusan különböző módon valósítható meg. Jelöléssel teljesen aszimmetrikussá tett kockával a szabályos kitöltés ennyi módon lehetséges. Két kitöltést akkor tekintünk azonosnak, ha van olyan a csúcs- él- és lapilleszkedést megtartó bijektív leképezés (kombinatorikus izomorfizmus), amely a megjelölt kockákkal való kitöltések egyikét a másikhoz rendeli. A számítás eredményeként mindegyik lehetőséghez megkaptuk a kikövezést végrehajtó csoport generátorait, és definiáló relációit. Az így megkapott csoportok mindegyike izomorf a 219 euklideszi kristálycsoport [2] valamelyikével. Hogy melyikkel, annak eldöntése további vizsgálatot igényelt. Eredményként azt kaptuk, hogy a 298 esetet összesen 130 kristálycsoport valósítja meg, tehát ennyi kristálycsoportnak lehet alaptartománya a kocka [8]. Ezek között természetesen vannak olyanok is, amelyekhez többféle módon válszthatunk kocka alaptartományt. Illusztrációként megadjuk azt a 45 esetet, 31 csoportot, ahol a kocka 6 lapját 3 generátor párosítja. Ezek közül 18 eset 9 fixpontmentes tércsoport hoz létre [4, 7].

1. Definíciók, elnevezések

Jelölje V , E és F rendre a poliéder csúcsainak, éleinek és lapjainak halmazát. Az egymáshoz illeszkedő csúcs, él, lap rendezett hármasok a poliéder zászlóit alkotják. A zászlók F halmaza tehát része a $V \times E \times F$ halmaznak. Feltesszük, hogy a csúcsok, élek és lapok bármelyike valamely zászlóhoz tartozik. Egy adott csúcst tartalmazó zászlók halmazát a csúcs zászlóinak nevezzük. Hasonlóan definiáljuk egy adott él illetve egy adott lap zászlóit. F -re a következő két feltételnek kell teljesülni.

1. Minden f zászlóhoz található pontosan egy f_v csúcsszomszéd, pontosan egy f_e élszomszéd és pontosan egy f_f lapszomszéd zászló. Két zászló akkor csúcsszomszédos, ha csúcskomponensük különböző, de él- és lapkomponensük azonos. Hasonlóan definiálhatók az él- és lapszomszédossági relációk is.
2. Bármely két különböző f_a és f_b zászló esetén képezhető a zászlókból olyan f_1, f_2, \dots, f_n véges sorozat, amelyre $f_1 = f_a$, $f_n = f_b$; és f_m csúcs-, él- vagy lapszomszédos f_{m+1} -gyel ($1 \leq m \leq n-1$); továbbá a sorozatban szereplő bármely zászló megfelelő komponenseként tartalmazza f_a és f_b közös komponenseit. E feltételek teljesülése esetén azt mondjuk, hogy a poliéder, illetve zászlóinak F halmaza, szigorúan összefüggő. A továbbiakban csak ilyen poliéderekről lesz szó.

Jelölje F_i és F_j rendre a poliéder l_i és l_j lapjának zászlóit. A $\gamma : F_i \rightarrow F_j$ bijektív leképezést izomorfizmusnak nevezzük, ha a csúcs- és élszomszédossági relációkat megtartja. Ha létezik ilyen γ leképezés, akkor az l_i és l_j lapokat izomorfoknak mondjuk. Az izomorf lapokat egy osztályba sorolva a lapok izomorfiaosztályaihoz jutunk. (Háromdimenziós poliéder esetén pontosan azok a lapok tartoznak egy izomorfiaosztályba, amelyeknek az oldalszáma azonos.) Ha $l_i = l_j$ ugyanaz a lap, akkor γ az l_i lap automorfizmusa. Az egy izomorfiaosztályba tartozó lapok automorfizmus csoportjait izomorfak.

A $\beta : F \rightarrow F$ bijektív leképezést a poliéder lapleképezésének nevezzük, ha bármely lap zászlóinak halmazára megszorítva, a lap egy izomorf leképezéséhez jutunk. A lapleképezések csoportjának involutív elemeit lapazonosításnak vagy lappárosításnak nevezzük. Azokat a lapleképezéseket, amelyek a csúcs- és élszomszédossági relációkon kívül a lapszomszédosságokat is megtartják, a poliéder automorfizmusainak mondjuk. Az automorfizmusok a lapleképezések csoportjának egy részcsoportját alkotják.

A $\gamma : F \rightarrow F$ és $\delta : F \rightarrow F$ bijektív leképezéseket ekvivalensnek mondjuk, ha van olyan α automorfizmusa a poliédernek, hogy egy tetszőleges f_a zászlóra, ha $f_a^\gamma = f_b$, akkor $(f_a^\delta)^\delta = f_b^\delta = (f_a^\gamma)^\alpha$, vagyis $\delta = \alpha^{-1}\gamma\alpha$.

2. Generátorok és lapazonosítások

Ha egy tércsoportnak egy adott poliéder az alaptartománya, akkor a poliédert a vele szomszédos példányokra képező transzformációk generálják a csoportot, annak egy generátorrendszerét alkotják. Legyen g egy ilyen transzformáció, amely a poliédert a g jelű lapja mentén vele szomszédos példányra képezi úgy, hogy a leképezés során a g^{-1} jelű lap a g lapra képeződik. A leképezésnél g^{-1} zászlóit egy izomorfizmus képezi rá g zászlóira. Ez az izomorfizmus és a g transzformáció kölcsönösen egyértelműen meghatározzák egymást. A g^{-1} transzformáció a poliédert a g^{-1} lapja mentén vele szomszédos példányra képezi, és közben a g lap a g^{-1} lapra képeződik. A zászlók leképezését az előbbi izomorfizmus inverze adja.

Előfordulhat, hogy valamely generáló transzformáció úgy képezi rá a poliédert valamely lapja mentén vele szomszédos példányra, hogy közben a lap önmagára képeződik. Ekkor a transzformáció inverze azonos a transzformációval. Ez azt jelenti, hogy a neki megfelelő izomorfizmus a lap involutív automorfizmusa.

Ezek szerint egy generátorrendszer megadása a következő módon történhet. Először rendezetlen párokat készítünk a poliéder lapjaiból a következő feltételek figyelembe vételével.

1. Két lap csak akkor lehet egy párban, ha ugyanabba az izomorfiaosztályba tartoznak. Az is lehet, hogy a pár két lapja azonos.
2. Minden lap egy és csak egy párban szerepel.

Másodszor az így elkészített párok mindegyikéhez hozzárendelünk egy izomorfizmust, amely az első lap zászlóit a második lap zászlóira képezi, továbbá hoz-

zárendeljük ezen izomorfizmus inverzét is, mely a második lap zászlóit képezi az első lap zászlóira. Ha a párban szereplő két lap azonos, akkor a lap egy involutív automorfizmusát rendeljük a párhoz.

Ha a párokhoz rendelt izomorfizmusokat mint leképezéseket egyesítjük, akkor a poliéder egy lapleképezéséhez jutunk. Könnyen belátható, hogy ez a lapleképezés involutív, tehát lapazonosítás. Egy generátorrendszer ilyen módon meghatároz egy lapazonosítást. Belátható az is, hogy ez a megfeleltetés kölcsönösen egyértelmű. Ezért a lehetséges generátorrendszerek megkeresése egyet jelent a lapazonosítások megkeresésével. Két generátorrendszer ekvivalens — ezeket nem különböztetjük meg — ha a nekik megfelelő lapazonosítások ekvivalensek, vagyis, ha a poliéder valamely automorfizmusánál (kombinatorikus szimmetriájánál) az egyik generátorrendszer a másikkra képeződik.

3. A Poincaré algoritmus

A tércsoport megadásához a generátorokon kívül meg kell határozni a csoport definiáló relációit is. Ezek a relációk a generátorok és a poliéder ismeretében a Poincaré algoritmus segítségével adódnak [3, 6, 9].

Az algoritmus a következő gondolaton alapul. A poliéder egy élet kiválasztva, az él mentén csatlakozó képpoliéder példányok az élet tartalmazó lapok mentén szomszédosak. Ezért az él körüljárása a következő módon írható le. A poliédert jelölje P ($P = P^1$). P -t a kiválasztott élhez illeszkedő g_1 jelű lapja mentén vele szomszédos példányra a g_1 generátor képezi, és így a P^{g_1} példányhoz jutunk. Ezt valamely, a kiválasztott élhez illeszkedő g_2 jelű lapja mentén vele szomszédos példányra a $g_1^{-1}g_2g_1$ transzformáció képezi, és így a $P^{g_2g_1}$ példányhoz jutunk. Ilyen módon folytatva a körüljárást, végül a $P^{g_n g_{n-1} \dots g_1}$ példányhoz jutunk, amely azonos P^1 -gyel. Így a $g_n g_{n-1} \dots g_1 = 1$ relációt kapjuk. Ezután egy olyan élt választunk ki, amely az előző reláció felírása során egyik példányon sem lépett fel a kiválasztott élhez illeszkedve. Ezen él körüljárásával egy újabb relációhoz juthatunk. Ezt az eljárást folytatjuk addig, amíg az élek mindegyike sorra nem kerül valamelyik reláció feírása során.

Az algoritmus ennek megfelelően a következő. Legyen adott a γ lapazonosítás. Válasszunk ki egy e élet az E halmazból, és legyen f_1 e -nek egyik zászlója. Ezek után legyen f_2 az f_1^γ lapszomszédja, f_3 az f_2^γ lapszomszédja, és így tovább, míg végül az f_n^γ lapszomszédja f_1 lesz. Az f_1, f_2, \dots, f_n zászlók g_1, g_2, \dots, g_n lapkomponenseihez rendelt generátorokból képezzük a $c := g_n \dots g_2 g_1$ szorzatot. Ezt a szorzatot ciklus-transzformációnak nevezzük. Egy egyelőre tetszőleges k pozitív egész számmal kapjuk a $c^k = 1$ definiáló relációt. A felsorolt zászlók

$$e_1 := e, e_2 := e_1^{g_1^{-1}}, \dots, e_n := e_{n-1}^{g_{n-1}^{-1}}, e_n^{g_n^{-1}} = e$$

élkomponensei az éleknek egy ekvivalenciaosztályát alkotják. Ezek az élek a

$$g_1, g_2 g_1, \dots, g_n \dots g_2 g_1$$

transzformációknál rendre a kiinduló e irányított élre képeződnek. (Előfordulhat, hogy a fenti irányított élek között $e_{i+1} = e_i$, ha $g_i^{-1} = g_i$ sítükrözés.) A k szám megválasztása a poliéder ezen élekhez tartozó lapszögeivel függ majd össze.

A következő ciklustranszformációhoz úgy juthatunk, ha keresünk egy olyan élet E -ben, amely az eddig megtalált ciklustranszformációk keresése közben felsorolt zászlók egyikében sem szerepelt. Ha van ilyen, akkor az előbbi e él helyett ebből kiindulva, új ciklustranszformációhoz jutunk, és ennek pozitív egész kitevős hatványaként újabb definiáló relációt kapunk. Ellenkező esetben nincs több ciklustranszformáció. Ha valamely g_i laphoz tartozó generátor a lapot önmagára képezi, akkor a definiáló relációk közé fel kell venni a $g_i^2 = 1$ -et is. (Lásd még [1, 3].)

Ilyen módon egy \mathcal{G} csoporthoz jutunk. Ha a ciklus-transzformációk kitevőjét a definiáló relációk felírásakor tetszőlegesen választjuk, akkor általában \mathcal{G} nem lesz diszkrét csoportja egyetlen állandó görbületű térnek sem. További vizsgálatot igényel a realizálhatóság kérdése. (A tetraéder esetén pl. lásd [5]-ben.)

A poliéder és a \mathcal{G} csoport ismeretében kombinatorikusan megadható az a pont-halmaz (absztrakt tér), amelyben \mathcal{G} úgy hat, hogy a poliéder a csoport alaptartománya legyen. Ez a következő módon történik. Tegyük fel, hogy a poliéder, a poliéder lapjai illetve a poliéder élei rendre nyílt tetraéderrel, nyílt háromszöggel illetve nyílt szakasszal homeomorf pont-halmazok, és bármely g_i generátor a g_i^{-1} lapot g_i -re képező homeomorfizmust származtatja. Megszerkesztjük a

$$P^{\mathcal{G}} = \bigcup_{h \in \mathcal{G}} (P, h)$$

poliéder-együttest mint pont-halmazt, amelynek elemei $x^h := (x, h)$ pontok $\langle x, h \rangle$ ekvivalenciaosztályai lesznek. A \mathcal{G} transzformáció-csoport majd egyszerűen tranzitív módon hat a $P^{\mathcal{G}}$ poliéder-együttesen, ha $h, h_1 \in \mathcal{G}$ esetén $h : \langle x, h_1 \rangle \rightarrow \langle x, h_1 h \rangle$ értelmezi a megfelelő ponttranszformációt, ami homeomorfizmus lesz.

Legyenek tehát (x_a, h_a) és (x_b, h_b) ekvivalensek, ha a következő két feltétel valamelyike teljesül.

1. $(x_a, h_a) = (x_b, h_b)$, azaz $x_a = x_b$ és $h_a = h_b$
2. Létezik olyan

$$(x_1, h_1) = (x_a, h_a), (x_2, h_2), \dots, (x_r, h_r) = (x_b, h_b)$$

véges sorozat, hogy tetszőleges j -re $(1 \leq j \leq r-1)$ található olyan g_j generátor, amelyre teljesül, hogy

$$x_j \in g_j^{-1}, \quad x_{j+1} = x_j^{g_j} \in g_j \quad \text{és} \quad h_j = g_j h_{j+1}.$$

Ez azt jelenti, hogy a P^h példány g_i^{-1} lapjának pontjait illetve zászlóit azonosítjuk a P^h példányhoz a g_i^{-1} lapja mentén illeszkedő $P^{g_i^{-1}h}$ példány g_i lapjának megfelelő pontjaival illetve zászlóival.

4. A lehetséges generátorrendszerek megkeresése

Ismertetjük azt az algoritmust, amelynek segítségével a lapazonosítások halmaza ekvivalenciaosztályainak mindegyikéből kiválasztható egy-egy reprezentáns elem. Ezek meghatározzák a lehetséges lényegesen különböző generátorrendszereket, amelyekből a Poincaré algoritmussal adódnak a lehetséges csoportok.

Jelölje B az adott poliéder lapleképezéseinek csoportját, A az automorfizmusok részcsoportját, T pedig a lapazonosítások (involutív lapleképezések) részhalmazát. Mivel a lapleképezések a zászlók véges halmazának speciális permutációi, ezért a definíciók alapján ezek a halmazok elvileg egyszerűen létrehozhatók. Legyen továbbá Q kezdetben üres halmaz, amely az algoritmus végrehajtása után a reprezentáns lapazonosításokat fogja tartalmazni. Az algoritmus ezek után a következő.

Vegyük sorra T elemeit, és legyen a soron lévő elem γ_i . Minden γ_i esetén vegyük sorra A elemeit: a soron lévő elem legyen α_j . Minden α_j esetén képezzük az $\alpha_j^{-1}\gamma_i\alpha_j$ szorzatot, és ha ezek egyike sem szerepel Q -ban, akkor γ_i -t vegyük hozzá Q elemeihez ($Q := Q \cup \{\gamma_i\}$). Ha T minden elemét sorra vettük, akkor Q egy teljes reprezentáns rendszert fog tartalmazni. Q elemeit sorra véve és ezekkel a Poincaré algoritmust végrehajtva, a ciklustranszformációkhoz tartozó kitevők meghatározásával megkapjuk a lehetséges csoportokat.

5. Az algoritmust módosítása

Tegyük fel, hogy a poliéder illeszkedési struktúráján kívül metrikus jellemzőit is megadjuk. A poliéderről feltételezzük most is, hogy szigorúan összefüggő. Ezen kívül a metrikus jellemzőkkel kapcsolatban két feltétel teljesülését követeljük meg:

1. A poliéder minden automorfizmusa a metrikusan adott poliédert önmagára képező egybevágóságot határoz meg.
2. A poliéder minden lapja szabályos sokszög, és az azonos oldalszámú lapok egybevágók.

Ez utóbbi kikötéssel a következőt érhetjük el. Ha δ az f_1 lap zászlóit az f_2 lap zászlóira képező izomorfizmus, akkor k -hoz egyértelműen hozzárendelhető a térnek az a h egybevágósága, amely az f_1 lap bármely zászlóját úgy képezi le f_2 egy zászlójára, mint δ , és közben a poliédert az f_2 lap mentén vele szomszédos példányra képezi.

Határozzuk meg a feltételeknek eleget tevő poliéder minden éléhez a hozzá illeszkedő két lap hajlásszögét. (Kettő, a feltételeknek eleget tevő poliéder illeszkedési struktúrája lehet azonos úgy, hogy a megfelelő lapszögek különbözőek. Például az euklideszi térben realizálódó kocka lapszögei 90° -osak, a hiperbolikus térben realizálódó, a teret egybevágó példányaival kitöltő azonos illeszkedési struktúrájú kocka lapszögei lehetnek 72° -osak, sőt 60° -osak, ha a csúcsok az abszolút alakzaton lévő ideális pontok.)

Legyen adott most a P poliéder egy γ lapazonosítása. Tegyük fel, hogy γ meghatározza a térnek egy olyan diszkrét csoportját, amelynek P alaptartománya. A Poincaré algoritmus végrehajtása során ciklustranszformációkat kapunk. Például az f_1, f_2, \dots, f_n zászlók felsorolása közben a $c = g_n \cdots g_2 g_1$ ciklustranszformációt kapjuk. A felsorolt zászlók élkomponenseiből az e_1, e_2, \dots, e_n élsorozat képezhető. Ez azt jelenti, hogy az e_1 él körüljárása során rendre ezek az élek képeződnek e_1 -re, és közben a megfelelő példányoknak ezen élekhez tartozó lapszögei kerülnek e_1 -nél sorra egymás mellé. Mivel feltevésünk szerint γ meghatároz egy P alaptartományú csoportot, ezért az e_1, e_2, \dots, e_n élekhez tartozó lapszögeket összeadva, olyan szöget kapunk, amelynek valamilyen k egész számú többszöröse a teljesszög. Ebből adódik a $c_k = 1$ definiáló reláció. Megfordítva belátható, hogy ha a γ lapazonosításra igaz az, hogy a Poincaré algoritmus végrehajtása során adódó összes ciklustranszformációnál a fenti módon képezett szögösszegek mindegyikének egész számú többszöröse a teljesszög, akkor a γ által meghatározott generátorok egy P alaptartományú csoportját generálják az adott térnek.

Ennek alapján a feltételeinknek eleget tevő P poliéderhez a következő algoritmus szerint határozhatjuk meg az adott tér P alaptartományú csoportjait generáló lapazonosítások ekvivalenciaosztályainak egy teljes reprezentáns rendszerét. (A jelölések az előző algoritmusnál használt jelölésekkel megegyeznek.)

Legyen Q kezdetben üres halmaz. vegyük sorra T elemeit, és legyen a soron lévő elem γ_i . Minden γ_i esetén hajtsuk végre a Poincaré algoritmust, és a fellépő ciklustranszformációknál képezzük az előzőekben leírt szögösszeget. Ha a szögösszegek mindegyikének van olyan egész számú többszöröse, ami éppen a teljesszög, akkor vegyük sorra A elemeit: a soronlévő elem legyen α_j . Minden α_j -vel képezzük az $\alpha_j^{-1} \gamma_i \alpha_j$ szorzatot, és ha ezek egyike sem szerepel Q -ban, akkor γ_i -t vegyük fel Q elemei közé. Miután T összes elemét soravettük, Q a keresett reprezentánsok halmaza lesz. Q elemeivel a Poincaré algoritmust végrehajtva, és a ciklustranszformációk kitevőjét az előzőek szerint választva a keresett diszkrét csoportok definiáló relációit is megkapjuk.

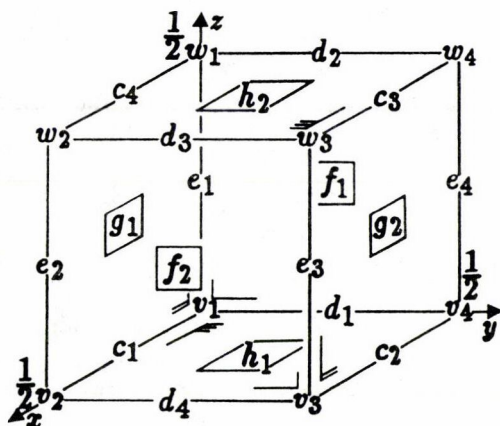
6. A kristálycsoportok visszakeresése a nemzetközi táblázatokból

Legyen most a vizsgált poliéder a kocka. (A továbbiakban az euklideszi teret röviden térnek, a benne realizálódó kockát röviden kockának mondjuk.) Az előbb leírt algoritmust erre alkalmazva megkaptuk a kristálycsoportot generáló lapazonosítások ekvivalenciaosztályainak egy teljes reprezentáns rendszerét, amit a továbbiakban a Q_{kocka} szimbólummal fogunk jelölni. A Q_{kocka} halmazban szereplő minden lapazonosításhoz megkaptuk az általa meghatározott generátorrendszert és a Poincaré algoritmus segítségével, a lapszögeket figyelembe véve, a generált kristálycsoport definiáló relációit.

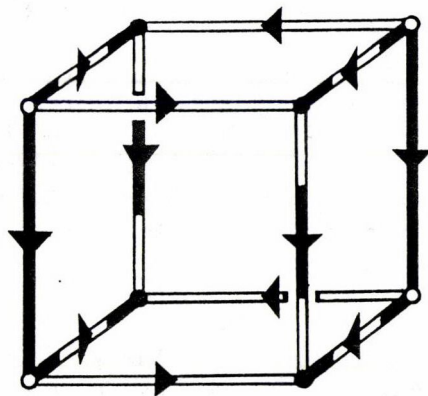
Annak meghatározása, hogy egy így megkapott kristálycsoport a tér 219 kristálycsoportja közül melyikkel izomorf, további vizsgálatot igényel. A vizsgálat alapja az, hogy a generáló transzformációk ismeretében meghatározzuk a vizsgált

csoporthoz tartozó tipikus transzformációkat, és így a [2] táblázatai alapján azonosítjuk a csoport nevét. Egy példán mutatjuk meg azt, hogy egy számítógép által megtalált csoportot hogyan keresünk vissza a [2] táblázataiból.

Egy feltételeinknek eleget tevő \mathcal{G} csoport a következő formában adódik az algoritmus végrehajtása során:



1. ábra



2. ábra

Generátorok:

$$\alpha : (v_1, d_1, f_1) \rightarrow (v_3, d_4, f_2) \quad [\text{---}]$$

$$\beta : (v_1, e_1, g_1) \rightarrow (v_3, e_3, g_2) \quad [\text{---}]$$

$$\gamma : (v_1, c_1, h_1) \rightarrow (w_3, c_3, h_2) \quad [\text{---}]$$

Definiáló relációk:

$$(v_1, d_1, f_1) \xrightarrow{\alpha} (v_3, d_4, f_2) \square\square (v_3, d_4, h_1)$$

$$(v_3, d_4, h_1) \xrightarrow{\gamma} (w_1, d_2, h_2) \square\square (w_1, d_2, f_1)$$

$$(w_1, d_2, f_1) \xrightarrow{\alpha} (w_3, d_3, f_2) \square\square (w_3, d_3, h_2)$$

$$(w_3, d_3, h_2) \xrightarrow{\gamma^{-1}} (v_1, d_1, h_1) \square\square (v_1, d_1, f_1)$$

$$\alpha\gamma\alpha^{-1} = 1$$

$$(v_1, e_1, g_1) \xrightarrow{\beta} (v_3, e_3, g_2) \square\square (v_3, e_3, f_2)$$

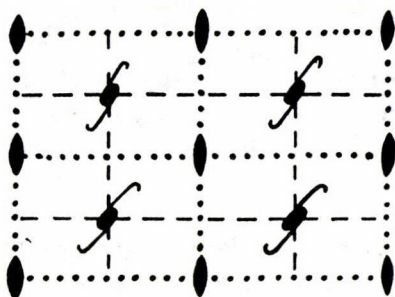
$$(v_3, e_3, f_2) \xrightarrow{\alpha^{-1}} (v_1, e_1, f_1) \square\square (v_1, e_1, g_1)$$

$$(\beta\alpha^{-1})^2 = 1$$

kiszámításával kapjuk a G/P faktorcsoporthoz elemeit:

1^*	$=$	(x, y, z)	egységelem
a_y^*	$=$	$(x + \frac{1}{2}, -y + \frac{1}{2}, z)$	eltolástükrözés (α)
b_x^*	$=$	$(-x + \frac{1}{2}, y + \frac{1}{2}, z)$	eltolástükrözés (β)
2_z^*	$= a_y^* b_x^* =$	$(-x, -y, z)$	félfordulat
2_{1z}^*	$=$	$(-x + \frac{1}{2}, -y + \frac{1}{2}, z + \frac{1}{2})$	csavarmozgás (γ)
c_x^*	$= a_y^* 2_{1z}^* =$	$(-x, y, z + \frac{1}{2})$	eltolástükrözés
c_y^*	$= b_x^* 2_{1z}^* =$	$(x, -y, z + \frac{1}{2})$	eltolástükrözés
i^*	$= a_y^* b_x^* 2_{1z}^* =$	$(x + \frac{1}{2}, y + \frac{1}{2}, z + \frac{1}{2})$	tércentráló eltolás.

A faktorcsoporthoz ismeretében elkészíthetjük a \mathcal{P} által meghatározott rács egy cellájának „térképét”, ami szintén segíti az azonosítást:



3. ábra

A faktorcsoporthoz fellépet az $(x + \frac{1}{2}, y + \frac{1}{2}, z + \frac{1}{2})$ eltolás által reprezentált i^* mellékosztály. A G -ben szereplő eltolások részcsoportja tehát bővebb \mathcal{P} -nél. Az eltolások részcsoportja az $\mathcal{I} = \{p_x, p_y, p_z, i\}$ lesz, ami szintén invariáns részcsoportja G -nek. A G/\mathcal{I} faktorcsoporthoz G/\mathcal{P} ismeretében egyszerűen adódik:

1^+	$=$	(x, y, z)	egységelem
a_y^+	$=$	$(x + \frac{1}{2}, -y + \frac{1}{2}, z)$	eltolástükrözés (α)
b_x^+	$=$	$(-x + \frac{1}{2}, y + \frac{1}{2}, z)$	eltolástükrözés (β)
2_z^+	$= a_y^+ b_x^+ =$	$(-x, -y, z)$	félfordulat.

(Itt már (x', y', z') jelentése $(x, y, z) \rightarrow (x' + t_x + \frac{1}{2}u, y' + t_y + \frac{1}{2}u, z' + t_z + \frac{1}{2}u)$, ahol t_x, t_y és t_z egészek, u pedig 0 vagy 1.) A keresett csoporthoz tehát tércentrált téglalánc tartozik.

Az előzőekben leírtakat figyelembe véve [2]-ben a 131. oldalon azonosíthatjuk a G csoportot: a 45-ös sorszámú $Iba2$ csoportot kapjuk.

7. Az eredmények összefoglalása

Eredményeinket két tételben foglalhatjuk össze. Az első tétel a Q_{kocka} halmaz számosságára vonatkozik. Ez ad választ arra a kérdésre, hogy a megjelölt kockával hányféle kombinatorikusan különböző módon követhető ki szabályosan a tér.

1. TÉTEL. Q_{kocka} elemeinek száma 298 [8].

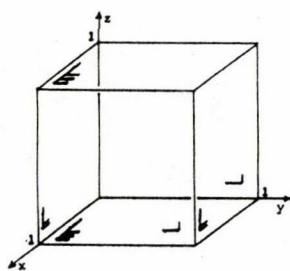
A 298 kombinatorikusan különböző kikövezés mindegyike valamely kristálycsoport transzfomációjánál valósul meg. Ezek a kristálycsoportok lesznek azok, amelyeknek alaptartománya a kocka. A második tételben ezeket a csoportokat adjuk meg.

2. TÉTEL. A 219 egymással páronként nem izomorf kristálycsoport közül az 1. táblázat felsorolt 130 csoportnak lehet a kocka alaptartománya. Minden csoport mellett megadjuk azt is, hogy a Q_{kocka} halmaznak hány eleme határozza meg az adott csoportot [8].

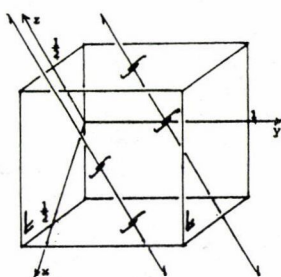
Végül kiemeljük Q_{kocka} azon elemeit, amelyek mint lapazonosítások, minden lap zászlóit egy másik lap zászlóira képezik. Az ilyen lapazonosításoknak megfelelő generátorrendszerek elemei a kocka minden lapját egy másik lapra képezik. Mivel a kockának hat lapja van, modhatjuk, hogy az így kapott csoportok lesznek a három elemmel generálható kocka alaptartományú kristálycsoportok. A generátorok száma itt geometriailag értendő, ugyanis lehet, hogy valamelyik generátor a másik kettővel kifejezhető. Q_{kocka} -nak 45 ilyen eleme van, ezek összesen 31 kristálycsoportot határoznak meg. Ezeket a 2. táblázat ábrasoraiban adjuk meg. A transzformációk jelölésére használt szimbólumokat a táblázat végén találhatjuk meg.

1. $P1$	(1)	34. $Pnn2$	(1)	68. $Ccca$	(4)	109. $I4_1md$	(1)
2. $P\bar{1}$	(3)	35. $Cmm2$	(1)	69. $Fmmm$	(1)	110. $I4_1cd$	(1)
3. $P2$	(2)	36. $Cmc2_1$	(3)	70. $Fddd$	(1)	112. $P\bar{4}2c$	(2)
4. $P2_1$	(2)	37. $Ccc2$	(1)	71. $Immm$	(1)	113. $P\bar{4}2_1m$	(1)
5. $B2$	(4)	38. $Amm2$	(1)	72. $Ibam$	(3)	114. $P\bar{4}2_1c$	(2)
6. Pm	(1)	39. $Abm2$	(2)	73. $Ibca$	(2)	115. $P4m2$	(1)
7. Pb	(3)	40. $Ama2$	(2)	74. $Imma$	(2)	116. $P\bar{4}c2$	(1)
8. Bm	(2)	41. $Aba2$	(4)	75. $P4$	(1)	117. $P\bar{4}b2$	(2)
9. Bb	(2)	42. $Fnm2$	(1)	76. $P4_1$	(1)	118. $P\bar{4}n2$	(1)
10. $P2/m$	(2)	44. $Imm2$	(1)	79. $I4$	(1)	119. $I\bar{4}m2$	(1)
11. $P2_1/m$	(3)	45. $Iba2$	(2)	80. $I4_1$	(1)	120. $I\bar{4}c2$	(2)
12. $B2/m$	(8)	46. $Ima2$	(2)	81. $P\bar{4}$	(1)	121. $I\bar{4}2m$	(1)
13. $P2/b$	(10)	47. $Pmmm$	(1)	82. $I\bar{4}$	(1)	122. $I\bar{4}2d$	(1)
14. $P2_1/b$	(13)	48. $Pnnn$	(1)	83. $P4/m$	(1)	124. $P4/mcc$	(1)
15. $B2/b$	(10)	49. $Pccm$	(2)	84. $P4_2/m$	(1)	125. $P4/nbm$	(1)
16. $P222$	(1)	50. $Pban$	(3)	85. $P4/n$	(2)	126. $P4/nnc$	(2)
17. $P222_1$	(3)	51. $Pmma$	(2)	86. $P4_2/n$	(1)	127. $P4/mbm$	(1)
18. $P2_12_12$	(4)	52. $Pnna$	(6)	87. $I4/m$	(2)	128. $P4/mnc$	(2)
19. $P2_12_12_1$	(1)	53. $Pmna$	(4)	88. $I4_1/a$	(2)	130. $P4/ncc$	(3)
20. $C222_1$	(4)	54. $Pcca$	(7)	89. $P422$	(1)	131. $P4_2/mmc$	(1)
21. $C222$	(2)	55. $Pbam$	(3)	90. $P42_12$	(2)	133. $P4_2/nbc$	(2)
22. $F222$	(1)	56. $Pccn$	(4)	91. $P4_122$	(2)	134. $P4_2/nnm$	(1)
23. $I222$	(2)	57. $Pbcm$	(5)	92. $P4_12_12$	(3)	135. $P4_2/mbc$	(2)
24. $I2_12_12_1$	(2)	58. $Pnnm$	(3)	93. $P4_222$	(1)	136. $P4_2/mnm$	(1)
25. $Pmm2$	(1)	59. $Pmmn$	(2)	94. $P4_22_12$	(1)	137. $P4_2/nmc$	(1)
26. $Pmc2_1$	(2)	60. $Pbcn$	(9)	97. $I422$	(2)	138. $P4_2/ncm$	(2)
27. $Pcc2$	(1)	61. $Pbca$	(5)	98. $I4_122$	(1)	140. $I4/mcm$	(1)
28. $Pma2$	(2)	62. $Pnma$	(6)	100. $P4bm$	(1)	141. $I4_1/amd$	(1)
29. $Pca2_1$	(4)	63. $Cmcm$	(4)	103. $P4cc$	(1)	142. $I4_1/acd$	(2)
30. $Pnc2$	(3)	64. $Cmca$	(7)	104. $P4nc$	(1)	144. $P3_1$	(1)
31. $Pmn2_1$	(2)	65. $Cmmm$	(1)	105. $P4_2mc$	(1)	152. $P3_121$	(1)
32. $Pba2$	(3)	66. $Cccm$	(2)	106. $P4_2bc$	(1)		
33. $Pna2_1$	(3)	67. $Cmma$	(2)	108. $I4cm$	(1)		

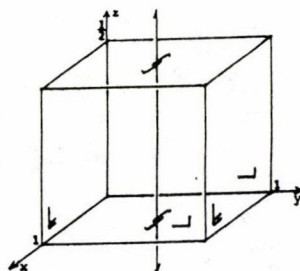
1. táblázat



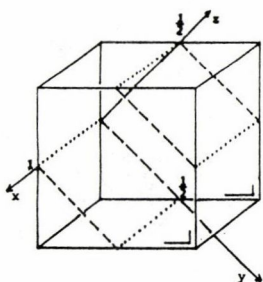
1.P1



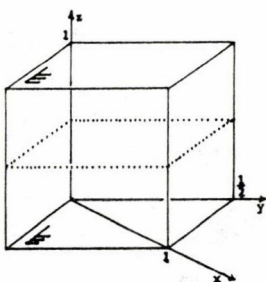
4.P2₁



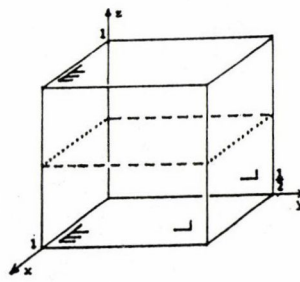
4.P2₁



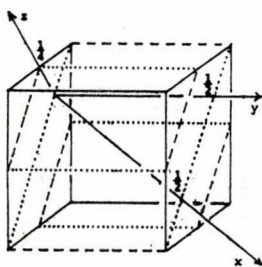
7.Pb



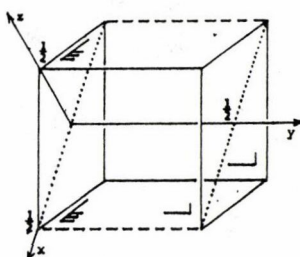
7.Pb



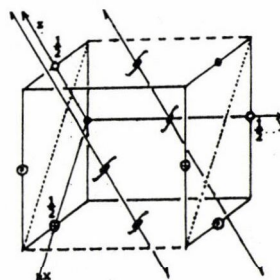
7.Pb



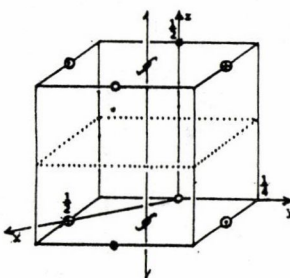
9.Bb



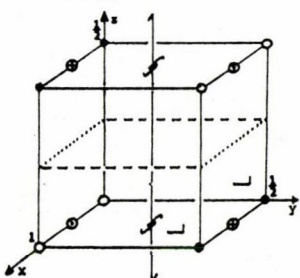
9.Bb



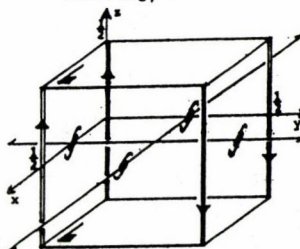
14.P2₁/b



14.P2₁/b

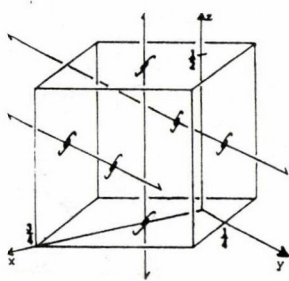
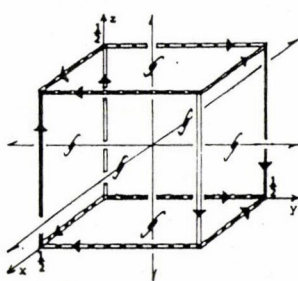
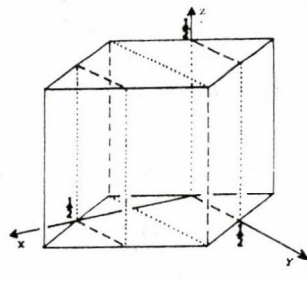
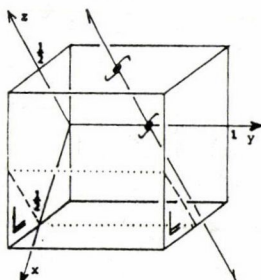
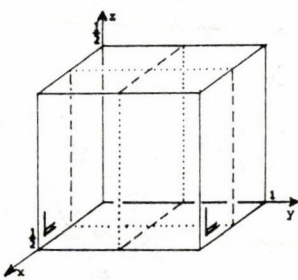
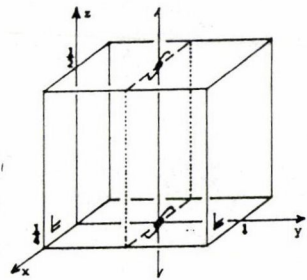
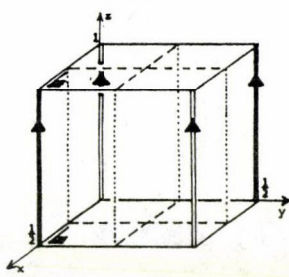
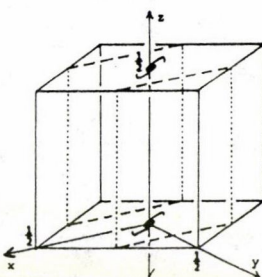
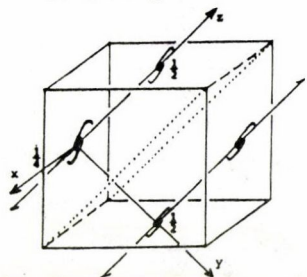
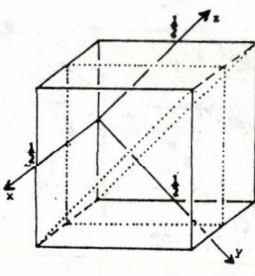
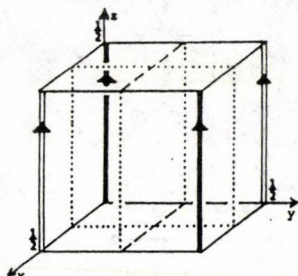
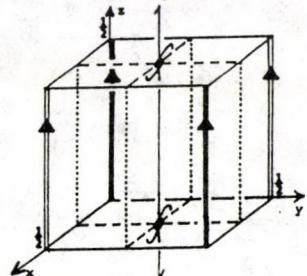


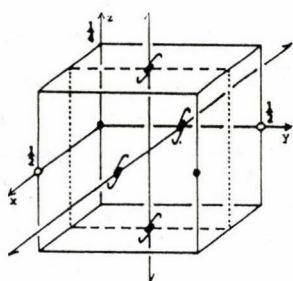
14.P2₁/b



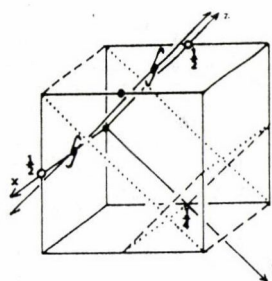
18.P2₁2₁2

2. táblázat (1. ábrásor)

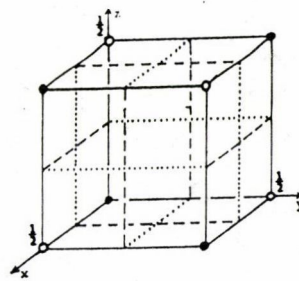
**19.P₂₁2₁2₁****23.I222****29.Pca2₁****29.Pca2₁****29.Pca2₁****29.Pca2₁****32.Pba2****33.Pna2₁****33.Pna2₁****33.Pna2₁****41.Aba2****45.Iba2****2. táblázat (2. ábrásor)**



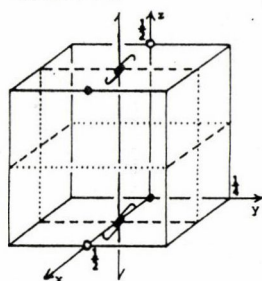
60.Pbcn



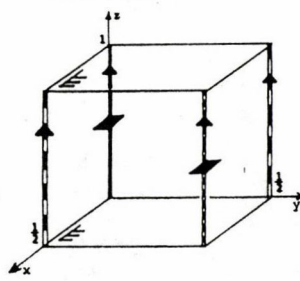
61.Pbca



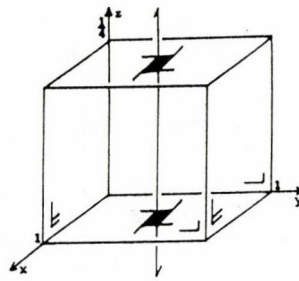
61.Pbca



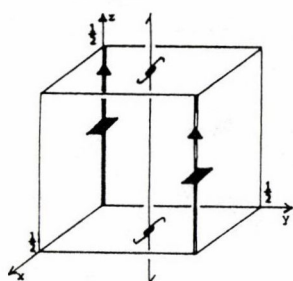
61.Pbca



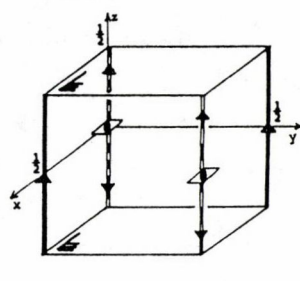
74.P4



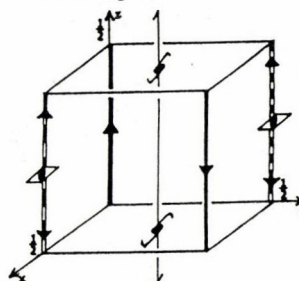
76.P4₁



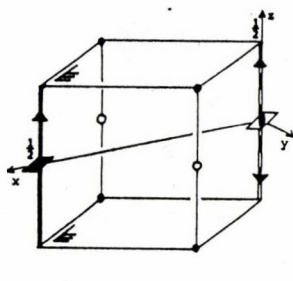
79.I4



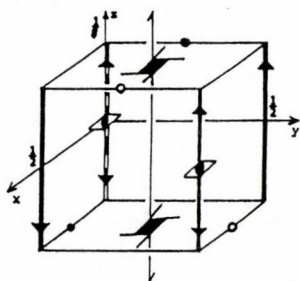
81.P4



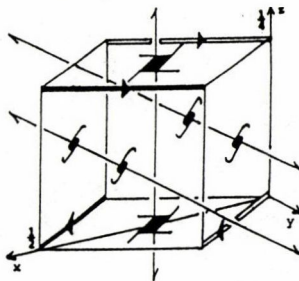
82.I4



85.P4/n

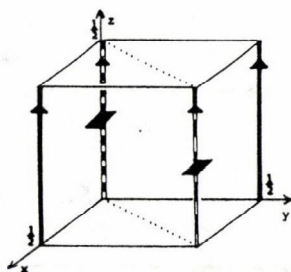
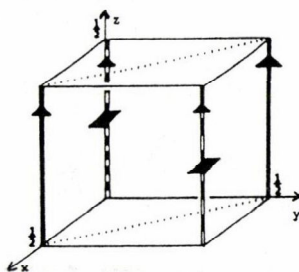
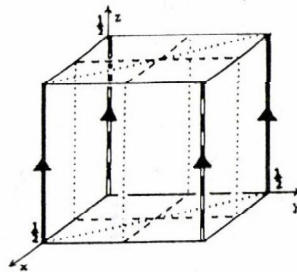
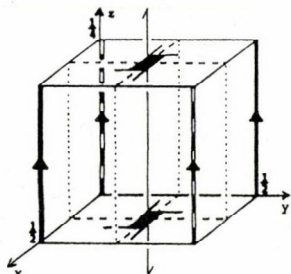
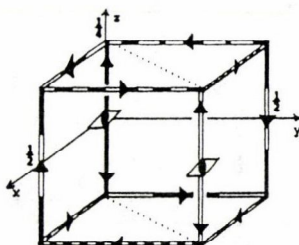
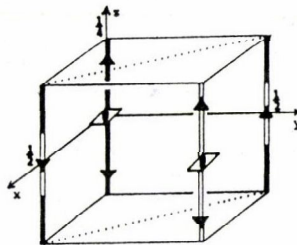
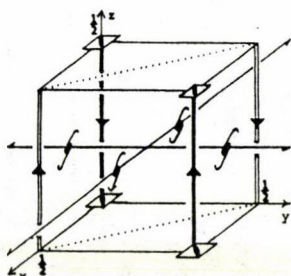
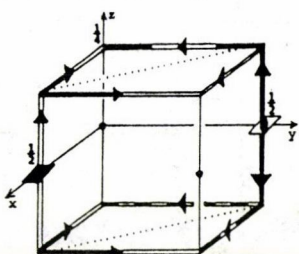
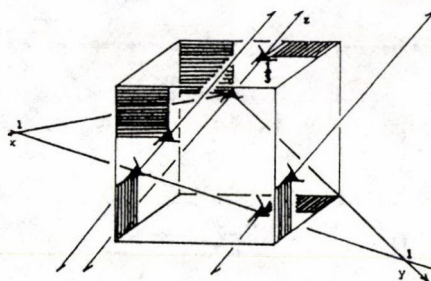


88.I4₁/a



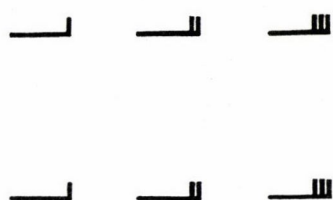
92.P4₁2₁2

2. táblázat (3. ábrásor)

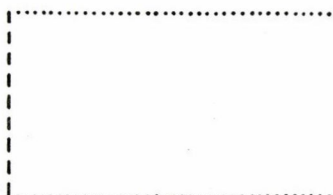
103. $P4_{cc}$ 104. $P4_{nc}$ 106. $P4_2bc$ 110. $I4_1cd$ 112. $P4_2c$ 114. $P4_2c_1$ 114. $P4_2c_1$ 130. $P4/ncc$ 145. $P3_1$

2. táblázat (4. ábrásor)

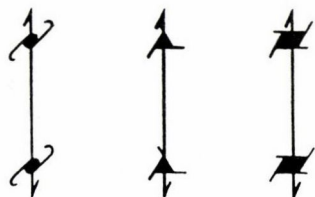
2. táblázat (jelölések)



Eltolás: az eltolást jellemző v vektor az ábra szerint adódik.



Eltolástükrözés: a jelölés a tükrözés síkjára utal; az eltolási komponens az ábra szerinti v vektor határozza meg.



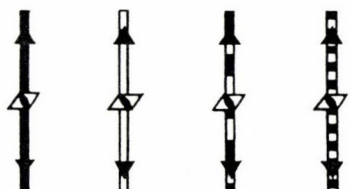
Csavarmozgás: a forgástengelyt a „fél”-nyilakkal ellátott szakaszok jelölik; az elforgatás szögére a tengelyen látható szimbólumok utalnak (az ábrán rendre 180° , 120° és 90° az elforgatás szöge); az eltolási komponens az ábra szerinti v vektor határozza meg.



Pontra vonatkozó tükrözés: a különböző jelek a centrumok különböző ekvivalenciaosztályaira utalnak.



Tengely körüli forgatás: széles vonal jelöli a tengelyt; (a vonalra rajzolt nyíl és a vonalak különféle mintái a megfelelő élek különböző ekvivalenciaosztályaira utalnak); a tengelyre rajzolt tömör négy-szög 90° -os forgatásra utal, egyébként a forgatás szöge 180° .



Forgatástükrözés: széles vonalak jelölik a tengelyt; a négy-szög a tengelyre merőleges tükörsík metszéspontját jelöli, továbbá azt, hogy a forgatás szöge 90° (így egy 180° -os forgatás is fellép az adott tengely körül); (a vonalak különféle mintái a megfelelő élek különböző osztályaira utalnak).

A tércsoportok felsorolásakor a csavarmozgást tartalmazó csoportok esetében megkülönböztetik azokat az egyébként izomorf csoportokat, amelyekben a csavarmozgások orientációja különböző. Így a tércsoportok száma 230. Ennek megfelelően

a kocka alaptartományú csoportok száma is megváltozik, 135 ilyen csoport van. Az ilyen módon adódó 5 új csoport a következő

78. $P_{4_3} \sim 76. P_{4_1}$
 96. $P_{4_3 2_1 2} \sim 92. P_{4_1 2_1 2}$
 95. $P_{4_3 2_2} \sim 91. P_{4_1 2_2}$
 145. $P_{3_2} \sim 144. P_{3_1}$
 154. $P_{3_2 2_1} \sim 152. P_{3_1 2_1}$

IRODALOM

- [1] COXETER, H. S. M. and MOSER, W. O. J., *Generators and relations for discrete groups*, 4th edition, *Ergeb. der Math.*, Neue Folge, Bd. 14., (Springer-Verlag, Berlin-Heidelberg-New York, 1980).
- [2] *International Tables for X-Ray Crystallography*, Vol. I.: Henry, N. F. M. – Lonsdale, K.; *Symmetry groups* New edition: Vol. A., edited by Theo Hahn, (Reidel, Birmingham, 1983).
- [3] MASKIT, B., „On Poincaré's theorem for fundamental polygons”, *Adv. in Math.* 7 (1971), 219–230.
- [4] MOLNÁR, E., „Minimal presentation of the 10 compact Euclidean space forms by fundamental domains”, *Studia Sci. Math. Hung.* 22 (1987), 19–51.
- [5] MOLNÁR, E. and PROK, I., „A polyhedron algorithm for finding space groups”, in *Proceedings of Third Int. Conf. on Engineering Graphics and Descriptive Geometry*, vol. 2 (Vienna, 1988), 37–44.
- [6] MOLNÁR, E., „Polyhedron complexes with simply transitive group actions and their realizations”, *Acta Math. Hung.* 59 (1992), 173–214.
- [7] PROK, I., „Térscsoportok keresése számítógéppel”, *diákköri dolgozat* (ELTE TTK Geometriai Tanszék, 1988).
- [8] PROK, I., „The Euclidean space has 298 fundamental tilings with marked cubes by 130 space groups.”, *Colloquia Math. Soc. János Bolyai* 63 *Intuitive Geometry*, Szeged (Hungary) 1991 (North-Holland Co., Amsterdam – Oxford – New York) (to appear).
- [9] VINBERG, E. B. and SHVARTSMAN, O. V., „Discrete transformation groups of spaces of constant curvature”, VINITI, *Itogi Nauki i Techniki, Sovr. Probl. Math. Fund. Napr.* 29 (1988), 147–259, (in Russian).

(Beérkezett: 1992. július 9.)

PROK ISTVÁN
 BME GEOMETRIA TANSZÉK
 BUDAPEST XI.
 EGRY JÓZSEF U. 1., H. II. 22
 1521

EUCLIDEAN CRYSTALLOGRAPHIC GROUPS WITH CUBIC FUNDAMENTAL DOMAIN I. PROK

The algorithm, described first in [5, 6], yields all the cubic space tilings where a marked cube is a fundamental domain of a crystallographic space group. The computer implementation has provided 298 combinatorially different fundamental tilings with marked cubes by 130 space groups listed in Table 3 [8]. In this paper 45 cubic tilings are described and illustrated where the six faces of the cube are paired by 3 generators. These represent 31 space groups.

NEMLINEÁRIS DIFFERENCIAEGYENLETEK GLOBÁLIS ASZIMPTOTIKUS STABILITÁSÁNAK ÁLTALÁNOS KRITÉRIUMAIRÓL

MOLNÁR SÁNDOR, SZIDAROVSKY FERENC, OKUGUCHI KOJI

Budapest, Budapest, Tokió

Általános kritériumokat adunk differenciaegyenletek globális aszimptotikus stabilitására, ahol az iterációs függvény halmazértékű is lehet, és még a pontértékű esetben sem kell feltenni a differenciálhatóságot. Megmutatjuk azt is, hogy az irodalomból ismert számos eredmény következik ezekből a feltételekből.

1. Bevezetés

Az elmúlt néhány évben egyre többen vizsgálták a nemlineáris differenciaegyenletek globális aszimptotikus stabilitását. BROCK és SCHEINKMAN (1975) a *Ljapunov függvény* deriváltja segítségével elégséges feltételt adott a globális stabilitásra, az iterációs függvény *Jacobi mátrixával* pedig egy speciálisabb elégséges feltételt nyertek. FUJIMOTO (1987a) ugyanennek az eredménynek egy némileg általánosabb feltételét fogalmazta meg, amikor az euklideszi normát tetszőleges vektornormával helyettesíthetjük. Ezt a tételt tovább lehetett általánosítani, mivel megmutatható, hogy elegendő, ha az iterációs függvény egy iteráltjára tesszük ugyanezeket a feltételeket (SZIDAROVSKY és OKUGUCHI 1988). FUJIMOTO (1987b) egy ehhez kapcsolódó cikkében HARTMAN (1961) jól ismert stabilitási feltételét adja meg a diszkrét idejű esetre.

A fenti eredmények esetében mindig feltették az iterációs függvény differenciálhatóságát, amely néhány alkalmazás esetén túlságosan erős feltevés. Ráadásul azt is feltételezik, hogy a differenciaegyenlet egy pont-pont leképezéssel adott. Ez is elég erős feltétel, mivel jónéhány alkalmazás esetén az iterációt egy pont-halmaz leképezés határozza meg. Cikkünkben egy pont-halmaz iterációval adott folytonos nemlineáris differenciaegyenletet tárgyalunk, és általános feltételt adunk a globális aszimptotikus stabilitásra. Alkalmazásképpen néhány korábbi eredményt általánosítunk. Végül fő eredményünket magasabb rendű differenciaegyenletekre is általánosítjuk.

2. Az általános modell

Legyen $S \subseteq R^n$ egy adott zárt halmaz, $0 \in S$, és legyen az f pont-halmaz leképezés az S -en definiálva úgy, hogy minden $x \in S$ esetén $\emptyset \neq f(x) \subseteq S$. Legyen

$x_0 \in S$ tetszőleges, és a rekurziót definiálja az

$$(1) \quad x_{k+1} \in f(x_k) \quad k = 0, 1, 2, \dots,$$

ahol az $f(x_k)$ halmazból tetszőlegesen választhatjuk ki az x_k utáni elemet. Tegyük fel, hogy fennállnak az alábbi feltételek:

- (A) az f leképezés felülről félig folytonos, azaz, ha $x_n \in S$ ($n \geq 0$), $x_n \rightarrow x^*$ és $n \geq 0$ -ra $y_n \in f(x_n)$ úgy, hogy $y_n \rightarrow y^*$, akkor $y^* \in f(x^*)$;
- (B) $f(0) = \{0\}$;
- (C) a $V : S \rightarrow R_+$ folytonos függvény;
- (D) $x \neq 0$ -ra és $y \in f(x)$ -re $V(y) < V(x)$;
- (E) tetszőleges (1) iterációs sorozathoz létezik egy $K > 0$ konstans úgy, hogy $\|x_k\| \leq K$ minden $k \geq 0$.

Fő eredményünket a következő tételben fogalmazzuk meg.

1. TÉTEL. Az (A)–(E) feltételek mellett az (1) sorozat nullához konvergál, azaz az origó globálisan, aszimptotikusan stabil az S halmazon.

Bizonyítás. Figyeljük meg először, hogy ha $x_k = 0$ valamely k -ra, akkor a (B) feltevésből következik, hogy minden $\ell \geq k$ esetén $x_\ell = 0$. Tegyük fel, hogy az x_k sorozat nem konvergál a nullához! Mivel a sorozat az (E) feltevés szerint egy kompakt halmazban van, létezik egy $\{x_{k_i}\}$ részsorozata, amely egy nem nulla x^* határértékhez konvergál. A (D) feltevésből, továbbá abból a tényből, hogy az x_k sorozatnak nincs nulla eleme, következik, hogy minden $i \geq 0$ -ra

$$(2) \quad V(x_{k_i+1}) \leq V(x_{k_i+1}) < V(x_{k_i}).$$

Mivel az $\{x_{k_i+1}\}$ is egy kompakt halmazban van, létezik konvergens részsorozata. Így feltehetjük, hogy az $\{x_{k_i}\}$ részsorozatot úgy választottuk meg, hogy az x_{k_i+1} sorozat is konvergens. Jelölje y^* ennek a második sorozatnak a határértékét! Ekkor az (A) feltevésből következik, hogy $y^* \in f(x^*)$. Ha a (2)-ben $i \rightarrow \infty$ a (C) feltevést felhasználva kapjuk, hogy $V(y^*) = V(x^*)$, ami ellentmond a (D) feltételnek. Ezzel a bizonyítást befejeztük.

Definiáljuk az f leképezés iteráltját a következőképpen: $f^1 \equiv f$ és $\ell \geq 1$ esetén

$$f^{\ell+1}(x) = \{y \mid y \in f(z) \quad z \in f^\ell(x) \text{ mellett}\}.$$

A következőkben az első tételt általánosítjuk.

2. TÉTEL. Tegyük fel, hogy az 1. tétel feltételei fennállnak úgy, hogy az f helyébe f^ℓ -et írunk, ahol az ℓ egy rögzített pozitív egész. Ekkor az origó globálisan aszimptotikusan stabil az S halmazon.

Bizonyítás. Az első tételből következik, hogy tetszőleges $z_0 \in S$ -re a

$$z_{k+1} \in f(z_k)$$

rekurzió által meghatározott sorozat a nullához konvergál. Belátjuk, hogy az (1) sorozatnak is a nullához kell konvergálnia. Ez az állítás rögtön következik abból, hogy a $z_0^{(j)} = x_j$ ($j = 0, 1, \dots, \ell - 1$) kezdeti vektorokra $z_k^{(j)} = x_{j+k\ell} \rightarrow 0$, ha $k \rightarrow \infty$ bármely rögzített j -re.

1. *Megjegyzés.* Az 1. tétel speciális esete a fenti eredménynek, ha az $\ell = 1$ választással élünk.

2. *Megjegyzés.* A 2. tétel esetén a (D) feltételt az alábbi szerint módosítjuk:

$$(D') \quad x \neq 0 \text{ és } y \in f^\ell(x)\text{-re } V(y) < V(x).$$

Figyeljük meg, hogy $(D) \Rightarrow (D')$, de $(D') \not\Rightarrow (D)$, ahogy azt az $f(x) = \{Ax\}$, $V(x) = \|x\|$ eset mutatja, ahol A egy olyan $n \times n$ -es konstans mátrix, amelynek minden sajátértéke az egységömbön belül van, de az $\|A\|$ nem kisebb az egységnél.

3. *Megjegyzés.* Megjegyezzük, hogy az

$$(E') \quad x \rightarrow \infty \text{-re } (x \in S) \quad V(x) \rightarrow \infty$$

feltételből következik az (E). Mégis sok szerző (pl. ZANGWILL 1969., BROCK és SCHEINKMANN 1975.) ezt az erősebb feltevést alkalmazza.

4. *Megjegyzés.* Tegyük fel, hogy S kompakt! Ekkor az (E) feltevésre nincs szükség. Tegyük továbbá fel, hogy az f pont-pont leképezés és folytonos!

Megjegyezzük, hogy ebben a speciális esetben az 1. tételt UZAWA (1961) stabilitási tétele diszkrét idejű megfelelőjének tekinthetjük.

Az 1. tétel alkalmazásakor a (D) feltétel teljesülése a legkritikusabb. A szakasz további részében elégséges feltételeket vezetünk be, amelyek garantálják (D) teljesülését. Az alábbi feltevéseket tesszük:

- (D1) Ha $x \in S$, akkor minden $0 \leq t \leq 1$ -re $tx \in S$;
- (D2) Minden $x \in S$ -re $f(x)$ korlátos. Mivel az (A) feltevésből következik, hogy $f(x)$ zárt, ezért az kompakt is.
- (D3) Létezik egy olyan $\varepsilon > 0$, hogy ha $\|x\| < \varepsilon$ és $x \in S$, akkor minden $y \in f(x)$ -re $V(y) - V(x) < 0$;
- (D4) Legyen $0 \leq t_1 < t_2 \leq 1$ és $x \in S$, továbbá $y_i \in f(t_i x)$ ($i = 1, 2$). Ekkor tetszőleges $V(y_1) - V(t_1 x)$ és $V(y_2) - V(t_2 x)$ közötti c számhoz létezik egy $\tau \in (t_1, t_2)$ úgy, hogy $V(y) - V(\tau x) = c$ valamely $y \in f(\tau x)$ -re;
- (D5) minden $0 \neq x \in S$ és $\varepsilon > 0$ -hoz létezik egy olyan $\delta > 0$, hogy ha $0 < h < \delta$ és $y_h \in f((1+h)x)$, akkor létezik egy olyan $y \in f(x)$, hogy $\|y - y_h\| < \varepsilon$;
- (D6) minden olyan $0 \neq x \in S$ -hez, amelyre valamilyen $y \in f(x)$, mellett $V(y) - V(x) = 0$ teljesül, létezik egy $0 < t < 1$ és egy $y_t \in f(tx)$, amelyekre $V(y_t) - V(tx) > 0$.

Ezek után megfogalmazhatjuk az alábbi következményt:

1. KÖVETKEZMÉNY. Az (A), (B), (C), (D1)–(D6) és (E) feltételek mellett az origó globálisan aszimptotikusan stabil.

Bizonyítás. Legyen az $x \neq 0, x \in S$ rögzített, és tekintsük a $T = \{t \mid t > 0 \text{ és valamely } y \in f(tx)\text{-re } V(y) - V(tx) \geq 0\}$ halmazt! Feltéve, hogy $T \neq \emptyset$, definiáljuk t_0 -at az alábbi módon:

$$t_0 = \inf\{t \mid t \in T\}.$$

A (D3) feltevésből következik, hogy $t_0 > 0$. Először belátjuk, hogy valamely $y \in f(t_0x)$ -re $V(y) - V(t_0x) = 0$. Tegyük fel először, hogy egy $y \in f(t_0x)$ -re $V(y) - V(t_0x) > 0$. A (D3) feltételből tudjuk, hogy kis $t > 0$ mellett $V(z) - V(tx) < 0$ minden $z \in f(tx)$ -re. Ekkor a (D4) feltételből következik, hogy létezik egy olyan $t^* \in (t, t_0)$, hogy $V(y^*) - V(t^*x) = 0$ valamely $y^* \in f(t^*x)$ -re, ami ellentmond a t_0 definíciójának. Tegyük most fel, hogy minden $y \in f(t_0x)$ -re $V(y) - V(t_0x) < 0$. Mivel az $f(t_0x)$ kompakt, és V folytonos, létezik egy olyan $\Delta > 0$, hogy minden $y \in f(t_0x)$ -re $V(y) - V(t_0x) < -\Delta$. Ekkor a (D5) feltevésből és a V függvény folytonosságából következik, hogy kis h mellett és $y_h \in f((1+h)t_0x)$ esetén $V(y_h) - V((1+h)t_0x) < 0$, ami ismét ellentmond a t_0 definíciójának. Ezért legalább egy $y_0 \in f(t_0x)$ -re $V(y_0) - V(t_0x) = 0$. Ekkor (D6) feltevésből következik, hogy létezik egy olyan $\tau < t_0$ ($\tau = (1-h)t_0x$), hogy $\tau \in T$, ami megint csak ellentmond a t_0 definíciójának. Ennélfogva a T -nek üres halmaznak kell lennie, azaz minden $t > 0$ -ra amelyre $tx \in S$, $V(y) < V(tx)$ minden $y \in f(tx)$ -szel. Mivel az $x \neq 0$ tetszőleges volt, az állítást beláttuk.

2. KÖVETKEZMÉNY. Tegyük most fel, hogy $f(x)$ minden $x \in S$ esetén egy pontból áll és folytonos az S halmazon! Ekkor teljesülnek az (A), (D2), (D4) és (D5) feltételek. Ezért a (B), (C), (E), (D1), (D3) és (D6) feltételek mellett az origó globálisan aszimptotikusan stabil.

Tegyük most fel, hogy az f halmazértékű, az S korlátos, és $x \in S$ -re definiáljuk y_x -et, amely a $V(y)$ függvényt maximalizálja az $y \in f(x)$ mellett!

Tegyük fel továbbá, hogy az y_x vektorok (több optimum esetén) megválaszthatók úgy, hogy az $f_0(x) = y_x$ függvény folytonos és kielégíti a fenti feltételeket! Ekkor minden $y \in f(x)$ -re

$$V(y) - V(x) \leq V(y_x) - V(x) = V(f_0(x)) - V(x) < 0.$$

Ekkor az 1. tétel (D) feltétele fennáll, így az origó globálisan aszimptotikusan stabil.

3. KÖVETKEZMÉNY. Tegyük fel, hogy az $f(x)$ csak egy pontból áll minden $x \in S$ -re, és $f(x) \in C^1$. Legyen $0 \neq x \in S$ olyan, hogy $V(f(x)) = V(x)$. Tegyük továbbá fel, hogy a $V \in C^1$ is teljesül! Ekkor a (D6) fennáll, ha a $V(f(x)) - V(x)$ x irány menti deriváltja az x pontban negatív. Mivel az iránymenti derivált

$$\lim_{\tau \rightarrow 0} \frac{1}{\tau} (f(x + \tau x) - f(x)),$$

kis negatív τ értékekre $f(x + \tau x) > f(x)$, ezért a (D6) fennáll.

A $V(f(x)) - V(x)$ x irány menti x pontbeli deriváltját a

$$\nabla V(f(x)) J(x) x - \nabla V(x) x$$

adja meg, ezért a (D6)-ot az garantálja, ha

$$V(f(x)) = V(x) \implies \nabla V(f(x)) J(x) x - \nabla V(x) x < 0.$$

5. *Megjegyzés.* Tegyük fel, hogy (D1) az S halmazra, a (C) feltétel a V függvényre teljesül, a (B) szerint az $f(0)$ egy pontból áll, és az (A), (D2), (D4) és (D5) folytonossági feltételek fennállnak az (E) mellett, akkor a maradék (D3) és (D6) feltételek szükségesek a (D)-hez. A (D3) azt a követelményt fogalmazza meg, hogy a (D)-beli egyenlőtlenség kis x értékekre fennáll. Ha ez minden $x \neq 0$ -ra fennáll, akkor nem létezik a $V(y) - V(x) = 0$ relációt kielégítő $x, y \in f(x)$ mellett, és ebből már következik az origó globális aszimptotikus stabilitása.

6. *Megjegyzés.* Helyettesítsük a (D3), (D5) és (D6) feltételeket az alábbiakkal: (D3') minden $0 \neq x \in S$ -hez létezik egy olyan $T > 0$, hogy $t \geq T$ és $y \in f(tx)$ esetén $V(y) - V(tx) < 0$;

(D5') minden $0 \neq x \in S$ és $\varepsilon > 0$ -hoz létezik egy olyan $\delta > 0$, hogy ha $0 < h < \delta$ és $y_h \in f((1-h)x)$, akkor létezik egy olyan $y \in f(x)$, hogy $\|y - y_h\| < \varepsilon$;

(D6') minden olyan $0 \neq x \in S$ -hez, amelyre valamilyen $y \in f(x)$ mellett $V(y) - V(x) = 0$ teljesül, létezik egy $t > 1$ és egy $y_t \in f(tx)$, amelyekre $V(y_t) - V(tx) > 0$. Ekkor az első következményben látottakhoz hasonlóan könnyen ellenőrizhető, hogy az (A), (B), (C), (D1), (D2), (D3'), (D4), (D5'), (D6') és (E) feltételekből következik a (D). Ráadásul ha a (D3') és (D6')-n kívül minden feltétel teljesül, akkor a (D3') és (D6') szintén szükséges feltételei annak, hogy a $V(y) - V(x) < 0$ reláció minden $0 \neq x \in S$ és $y \in f(x)$ -re teljesüljön.

7. *Megjegyzés.* Tegyük fel, hogy f pont-pont leképezés és az origó globálisan aszimptotikusan stabil! Ekkor létezik egy olyan $V : S \rightarrow \mathbb{R}_+$ függvény, hogy $x \neq 0$ -ra $V(f(x)) \leq V(x)$. Ha $x = 0$, akkor a (B) feltevésből következik, hogy $f(x) = 0$, és ha $x \neq 0$ definiáljuk az

$$x_0 = x, \quad x_{k+1} = f(x_k) \quad (k \geq 0)$$

sorozatot, továbbá legyen

$$V(x) = \max_{k \geq 0} \{\|x_k\|\}.$$

Ekkor nyilván

$$\begin{aligned} V(f(x)) &= V(x_1) \\ &= \max_{k \geq 1} \{\|x_k\|\} \\ &\leq \max_{k \geq 0} \{\|x_k\|\} \\ &= V(x). \end{aligned}$$

3. Speciális stabilitási feltételek

Tegyük most fel, hogy $f : R^n \rightarrow R^n$ egy olyan függvény, hogy $f(0) = 0$ és $f \in C^1$. Ekkor az $S = R^n$ -et választjuk. A V speciális megválasztása speciális stabilitási feltételekhez vezet.

Válasszuk először a $V(x) = \|x\|$ függvényt valamilyen vektornormával! A (D) feltétel most azt jelenti, hogy $x \neq 0$ -ra $\|f(x)\| < \|x\|$, és az 1. tétel összes többi feltétele fennáll. Ekkor az alábbi állítást fogalmazhatjuk meg.

1. ÁLLÍTÁS. Ha valamilyen norma mellett $\|f(x)\| < \|x\|$ minden $x \neq 0$ esetén, akkor az origó globálisan aszimptotikusan stabil

Megjegyzés. Ez FUJIMOTO (1987a) első állítása.

1. KÖVETKEZMÉNY. Az átlagértékre vonatkozó tételből (ORTEGA és RHEIN-BOLDT 1970) következik, hogy valamilyen norma mellett $\|J(x)\| < 1$ minden $x \neq 0$ -ra, akkor az origó globálisan aszimptotikusan stabil.

Megjegyezzük, hogy ez FUJIMOTO (1987a) második állítása.

2. KÖVETKEZMÉNY. Legyen most P egy konstans $n \times n$ -es nem szinguláris mátrix. Közismert, hogy ha $\|x\|$ valamilyen vektornormát jelöl, akkor $\|x\|_P = \|Px\|$ is vektornorma. Ezért, ha minden $x \neq 0$ -ra $\|Pf(x)\| < \|Px\|$, vagy $\|PJ(x)u\| < \|Pu\|$ (minden $u \neq 0$ -ra), akkor az origó globálisan aszimptotikusan stabil. Ha az $\|x\| = \sqrt{x^T x}$ normát választjuk, akkor ezek a feltételek ekvivalensek az

$$(3) \quad f^T(x)P^T Pf(x) < x^T P^T Px$$

és az

$$(4) \quad u^T (J(x)^T P^T P J(x) - P^T P) u < u^T u \quad (u \neq 0)$$

egyenlőtlenségekkel.

Megjegyezzük, hogy a (4) reláció minden $u \neq 0$ -ra akkor és csak akkor áll fenn, ha a

$$J(x)^T P^T P J(x) - P^T P$$

mátrix negatív definit. Ezt az eredményt kapta FUJIMOTO (1987b), amely OKUGUCHI (1976) 1.3.2.3. tételének általánosítása. Válasszuk most újra a $V(x) = \|x\|$ függvényt, akkor a 2. tétel alkalmazásával a következő állítást kapjuk.

2. ÁLLÍTÁS. Ha valamilyen norma mellett $\|f^\ell(x)\| < \|x\|$ minden $x \neq 0$ -ra, akkor az origó globálisan aszimptotikusan stabil.

3. KÖVETKEZMÉNY. Megjegyezzük, hogy az f^ℓ Jacobi mátrixa

$$J^\ell(x) = J(f^{\ell-1}(x)) \cdot J(f^{\ell-2}(x)) \dots J(f^1(x)) \cdot J(x).$$

Ezért, ha valamilyen pozitív egész ℓ számra $\|J^\ell(x)\| < 1$ minden $x \neq 0$ -ra, akkor az origó globálisan aszimptotikusan stabil.

Megjegyzés. Tekintsük az $f(x) = Ax$ lineáris esetet az A $n \times n$ -es konstans mátrix-szal! Ekkor $J^l(x) = A^l$, és így a 3. következmény feltétele szükséges és elégséges a globális aszimptotikus stabilitáshoz.

4. KÖVETKEZMÉNY. A fenti általánosított esetben a (3) feltételt az

$$f^l(x)^T P^T P f^l(x) < x^T P^T P x$$

egyenlőtlenséggel helyettesítettük, és a (4) feltétel ekvivalens volt azzal a feltevessel, hogy a

$$J^l(x)^T P^T P J^l(x) - P^T P$$

mátrix negatív definit legyen.

Tegyük fel újra, hogy a szakasz kezdetén megadott feltétel teljesül, továbbá azt, hogy $\lim_{\|x\| \rightarrow \infty} x^T f(x) = \infty$. Ha a $V(x) = x^T f(x)$ választással élünk, akkor az

1. tétel (D) feltétele azt jelenti, hogy

$$f^T(x) f(f(x)) < x^T f(x),$$

azaz

$$(5) \quad f^T(x) [f(f(x)) - x] < 0,$$

és a tétel összes többi feltétele érvényes. Így beláttuk a következő állítást:

3. ÁLLÍTÁS. Ha minden $x \neq 0$ -ra az (5) egyenlőtlenség fennáll, akkor az origó globálisan aszimptotikusan stabil.

A V bármely más lehetséges megválasztása hasonló módon tárgyalható, így a további részletektől eltekintünk.

4. Magasabbrendű differenciaegyenletek

Legyen S ismét az R^n egy zárt részhalmaza, és $0 \in S$. Tegyük fel, hogy az $f(t^{(1)}, t^{(2)}, \dots, t^{(\ell)})$ pont-halmaz leképezés definiálva van az $S^\ell = S \times S \times \dots \times S$ halmazon, és minden $(t^{(1)}, \dots, t^{(\ell)}) \in S$ esetén $\emptyset \neq f(t^{(1)}, \dots, t^{(\ell)}) \subseteq S$.

Legyenek $x_0, x_1, \dots, x_{\ell-1} \in S$ tetszőlegesek, és a rekurziót definiálja most az alábbi formula:

$$(6) \quad x_{k+1} \in f(x_{k-\ell+1}, \dots, x_k), \quad (k \geq \ell - 1)$$

ahol a képhalmazból tetszőlegesen választhatjuk meg az x_k utáni elemet. A következő feltevéseket tesszük:

- (A) Az f leképezés felülről félig folytonos, azaz $i = 1, 2, \dots$ esetén, ha az $\{x_n(i)\} \subset S$ konvergál az $x^{(i)*}$ -hoz, és minden $n \geq 0$ -ra $y_n \in f(x_n^{(1)}, \dots, x_n^{(i)})$ úgy, hogy $\{y_n\}$ az y^* -hoz konvergál, akkor $y^* = f(x^{(1)*}, \dots, x^{(i)*})$;
- (B) $f(t^{(1)}, \dots, t^{(\ell-1)}, 0) = 0$ minden $t^i \in S$ ($1 \leq i \leq \ell - 1$) esetén;
- (C) a $V : S \rightarrow R_+$ folytonos függvény;
- (D) $t^{(\ell)} \neq 0$ -ra és tetszőleges $t^{(i)} \in S$ ($1 \leq i \leq \ell - 1$) és $y \in f(t^{(1)}, \dots, t^{(\ell)})$ esetén $V(t^{(2)}, \dots, t^{(\ell)}, y) < V(t^{(1)}, \dots, t^{(\ell)})$;
- (E) tetszőleges (6) iterációs sorozatra létezik egy olyan $k > 0$, hogy $\|x_k\| \leq K$ (minden $k \geq 0$).

Általánosított eredményünket az alábbi tételben fogalmazzuk meg.

3. TÉTEL. Az (A)–(E) feltételek teljesülése esetén az $\{x_k\}$ sorozat nullához konvergál, azaz az origó globálisan aszimptotikusan stabil.

Bizonyítás. A (B) feltételből következik, hogy ha egy elemre $x_k = 0$, akkor az összes öt követő elem is nulla lesz. Tegyük most fel, hogy $x_k \rightarrow 0$. Ekkor $x_k \neq 0$ minden k -ra és létezik egy $\{x_{k_i}\}$ részsorozat, amely egy nem nulla x^* határértékhez konvergál, mivel az iterációs sorozat az (E) feltétel szerint egy kompakt halmazban van. Ekkor a (D) feltételből következik, hogy

$$(7) \quad V(x_{k_{i+1}-\ell+1}, \dots, x_{k_{i+1}}) \leq V(x_{k_i-\ell+2}, \dots, x_{k_{i+1}}) < V(x_{k_i-\ell+1}, \dots, x_{k_i}).$$

Az általánosság megszorítása nélkül feltehetjük, hogy minden más

$$\{x_{k_i-\ell+1}\}, \{x_{k_i-\ell+2}\}, \dots, \{x_{k_i-1}\} \text{ és } \{x_{k_{i+1}}\}$$

sorozat is konvergens. Jelölje ezek határértékeit rendre $x_{\ell-1}^*, x_{\ell-2}^*, \dots, x_1^*$ és y^* . Vegyük észre, hogy az (A) feltételből következően $y^* \in f(x_{\ell-1}^*, \dots, x_1^*, x^*)$, és ha a (7)-ben $i \rightarrow \infty$, akkor a

$$V(x_{\ell-2}^*, \dots, x_1^*, x^*, y^*) = V(x_{\ell-1}^*, \dots, x_1^*, x^*)$$

egyenlőséget kapjuk, amely ellentmond a (D) feltevésnek. Ezzel a bizonyítást befejeztük

Megjegyzés. A többlépcsős iterációs folyamatok fontos szerepet játszanak például az extrapolatív várakozásokkal jellemezhető dinamikus folyamatok stabilitásának vizsgálatában (lásd OKUGUCHI 1976 az oligopolista piacokról).

Az 1. tételhez hasonlóan a fenti eredményt is kiterjeszthetjük, illetve különböző speciális stabilitási feltételeket nyerhetünk.

IRODALOM

- [1975] BROCK, W. A., és SCHEINKMAN, J. A., „Some Results on Global Asymptotic Stability of Difference Equation”, *Journal of Economic Theory* 10 (1975), 265–268.
- [1987A] FUJIMOTO, T., „Global Asymptotic Stability of Non-Linear Difference Equations”, *Economics Letters* 22 (1987a), 247–250.
- [1987B] FUJIMOTO, T., „Global Asymptotic Stability of Non-Linear Difference Equations II”, *Economics Letters* 23 (1987b), 275–277.
- [1961] HARTMAN, P., „On Stability in the Large for Systems of Ordinary Differential Equations”, *Canadian Journal of Mathematics* 13 (1961), 480–492.
- [1976] OKUGUCHI, K., *Expectations and Stability in Oligopoly Models* (Springer-Verlag, Berlin/Heidelberg/New York, 1976).
- [1970] ORTEGA, J. M. and RHEINBOLD, W. C., *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).
- [1988] SZIDAROVSKY, F. and OKUGUCHI, K., „A Note on Global Asymptotic Stability of Non-linear Difference Equations”, *Economics Letters* 26 No.4 (1988), 349–352.
- [1961] UZAWA, H., „The Stability of Dynamic Processes”, *Econometrica* 29 No.4 (1961), 617–631.
- [1969] ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach* (Prentice-Hall, Englewood Cliffs, N.J., 1969).

(Beérkezett: 1992. április 25.)

MOLNÁR SÁNDOR
KÖZPONTI Bányászati Fejlesztési Intézet
BUDAPEST, MIKOVINY S. U. 2-4. 1037

SZIDAROVSKY FERENC
BUDAPESTI Közgazdaságtudományi Egyetem
MATEMATIKAI ÉS SZÁMÍTÁSTUDOMÁNYI INTÉZET
BUDAPEST, FÖVÁM TÉR 8. 1093

OKUGUCHI KOJI
DEPARTMENT OF ECONOMICS
TOKYO METROPOLITAN UNIVERSITY
1-1-1 YAKUMO MEGURO-KU, TOKYO, JAPAN

ON GENERAL CRITERIA FOR THE GLOBAL ASYMPTOTIC
STABILITY OF NONLINEAR DIFFERENCE EQUATIONS

S. MOLNÁR, F. SZIDAROVSKY and KOJI OKUGUCHI

General criteria for the global asymptotic stability of difference equations are derived, where the iteration functions may be set valued and even in the point valued case no differentiability is assumed. It is also shown that these new conditions imply several known results from the literature.

A külföldi szakirodalomból

MATEMATIKAI PROBLÉMÁK ÉS SEJTÉSEK NUMERIKUS MEGKÖZELÍTÉSBEN*

RICHARD S. VARGA

Kent

4. FEJEZET

Az $\exp(z)$ részletösszegei
zérushelyeinek aszimptotikája

4.1. Szegő tétele és a D_∞ és D_n görbék

Az e^z exponenciális függvény jólismert sorfejtését $s_n(z) := \sum_{j=0}^n z^j/j!$ ($n \geq 1$)

alakban írva, most az $s_n(nz)$ *normalizált* részletösszegek gyökeinek elhelyezkedését, valamint azt a sebességet vizsgáljuk, amellyel a gyökök a

$$(1.1) \quad D_\infty := \{z \in \mathbb{C} : |ze^{1-z}| = 1 \text{ és } |z| \leq 1\}$$

által definiált Szegő-görbéhez tartanak.

A jól ismert Eneström–Kakeya tétel szerint (lásd Marden [7, 137. oldal Ex. 2]) bármely $p_n(z) = \sum_{j=0}^n a_j z^j$ polinomra, ahol $a_j > 0$ ($0 \leq j \leq n$), igaz az, hogy zérushelyei a zárt

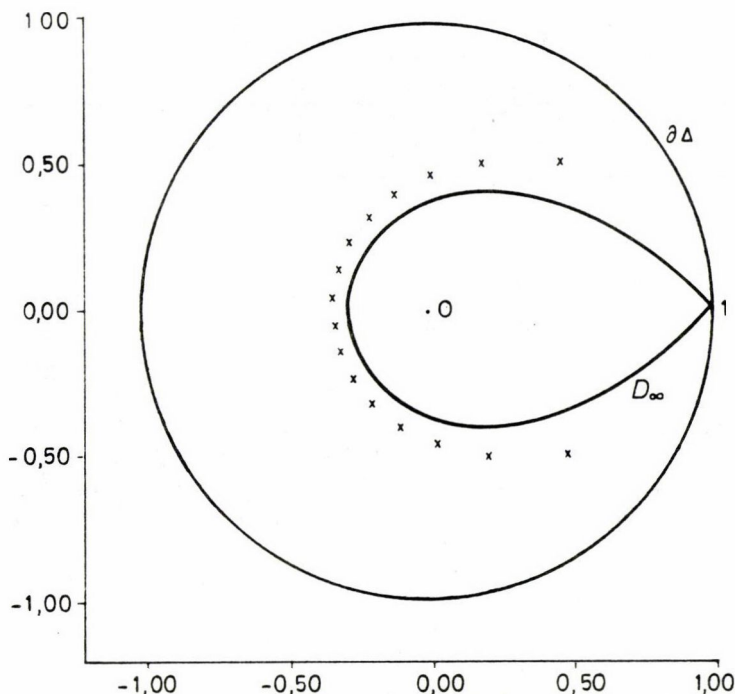
$$\min_{0 \leq i \leq n} \left(\frac{a_i}{a_{i+1}} \right) \leq |z| \leq \max_{0 \leq i < n} \left(\frac{a_i}{a_{i+1}} \right)$$

körgyűrűben fekszenek. Az utóbbi egyenlőtlenséget az e^z sorfejtésének $s_n(z)$ részletösszegeire alkalmazva azonnal kapjuk, hogy $s_n(z)$ valamennyi gyökére teljesül a $|z| \leq n$ egyenlőtlenség, bármely $n \geq 1$ esetén. Sőt, az Eneström–Kakeya tétel egy erősebb változata (lásd Anderson, Saff és Varga [1, Cor. 2]) azt állítja, hogy az $s_n(z)$ gyökei kielégítik a $|z| < n$ egyenlőtlenséget is, bármely $n > 1$ esetén. Így, ha $\{z_{k,n}\}_{k=1}^n$ jelöli az $s_n(nz)$ *normalizált* részletösszegek gyökeinek halmazát, akkor a gyökök a zárt $\Delta := \{z \in \mathbb{C} : |z| \leq 1\}$ egységkörlemezen fekszenek, bármely $n \geq 1$ esetén, továbbá a Δ *belsejében* vannak, bármely $n > 1$ mellett (ez világosan látszik a

*A fordítás RICHARD S. VARGA, Scientific Computation on Mathematical Problems and Conjectures című könyvéből a szerző és a kiadó engedélyével készült. Ez a rész a könyv negyedik, ötödik és hatodik fejezetének fordítása.

4.1. ábrán). Következésképpen, valamennyi normalizált $\{s_n(nz)\}_{n=1}^{\infty}$ részletösszeg összes gyökének végtelen halmaza *legalább* egy torlódási ponttal rendelkezik a Δ kompakt halmazban.

Szegő [12] nevezetes 1924-es cikkében bebizonyította a következőt:

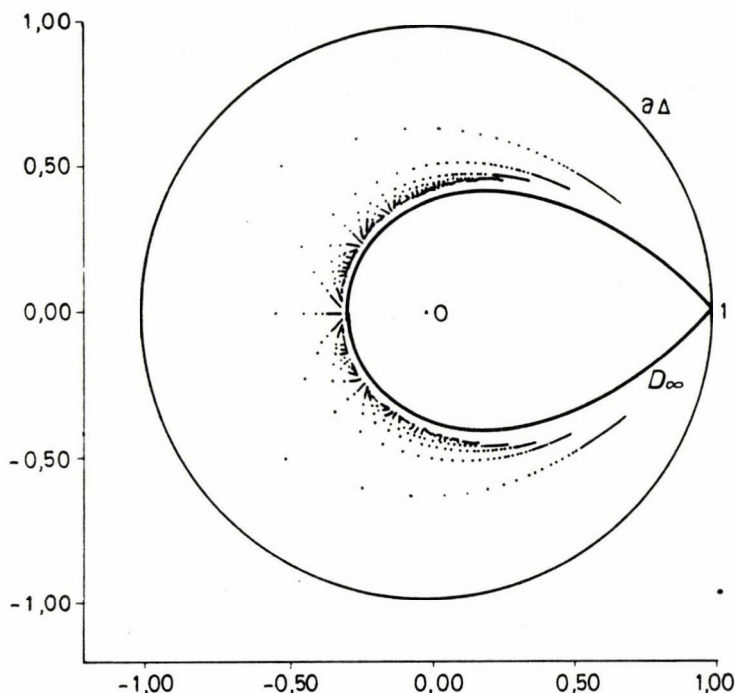


4.1. ábra: $\{s_n(nz)\}_{n=1}^{40}$ gyökei.

1. TÉTEL (Szegő [12]). Az e^z normalizált $\{s_n(nz)\}_{n=1}^{\infty}$ részletösszegeinek valamennyi $z \in \Delta$ torlódási pontja az (1.1) alatt definiált D_∞ görbén fekszik. Fordítva, a D_∞ bármely pontja egy torlódási pontja az e^z normalizált $\{s_n(nz)\}_{n=1}^{\infty}$ részletösszegei gyökeinek.

Később Buckholtz [2] megmutatta, hogy az $s_n(nz)$ gyökei a D_∞ görbén *kívül* fekszenek, bármely $n \geq 1$ esetén. Ezen eredmények szemléltetésére a 4.2 ábrán bemutattuk az $s_{16}(16z)$ függvény 16 gyökét (ezeket \times -ekkel ábrázolva), a D_∞ Szegőgörbével (lásd (1.1)) és $\partial\Delta$ -val, vagyis a Δ határvonalával együtt. Ugyanezt tettük a 4.3 ábrán az $s_{27}(27z)$ függvény 27 gyökével.

A 4.1, 4.2 és 4.3 ábrák jelzik, hogy az $s_n(nz)$ gyökei láthatóan „regulárisan” tartanak a D_∞ görbéhez, és hogy ez a konvergencia a *leglassabb* a D_∞ $z = 1$ pontjának környezetében. Tekintettel arra, hogy minket éppen e gyököknek a D_∞ görbéhez való *leglassabb* konvergenciájának sebessége érdekel, jelölje $\text{dist}[z, B] :=$



4.2 ábra: $s_{16}(16z)$ gyökei.

$\inf\{|z-w| : w \in B\}$ a szokásos (metrikus) távolságot a z pont és a B halmaz között a komplex síkon, és egy véges $\{t_j\}_{j=1}^m$ halmaz esetén legyen

$$(1.2) \quad \text{dist}[\{t_j\}_{j=1}^m; B] := \max_{1 \leq j \leq m} (\text{dist}[t_j, B]).$$

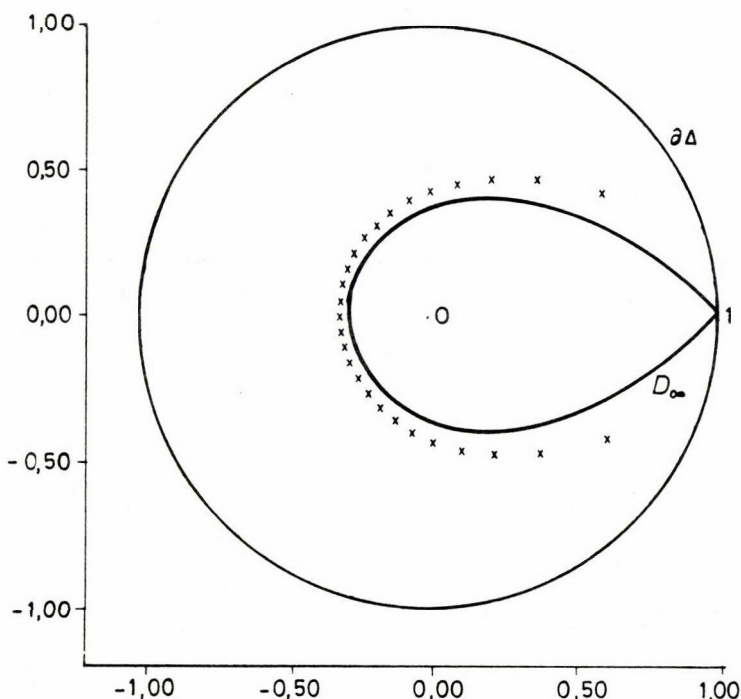
Buckholtz [2] bebizonyította, hogy az $s_n(nz)$ valamennyi $\{z_{k,n}\}_{k=1}^n$ gyöke legfeljebb $\frac{2e}{\sqrt{n}}$ távolságra van D_{∞} -tól, vagyis az (1.2) jelöléssel élve

$$(1.3) \quad \text{dist}[\{z_{k,n}\}_{k=1}^n; D_{\infty}] \leq \frac{2e}{\sqrt{n}} \quad (n = 1, 2, \dots).$$

Ebből természetesen következik, hogy

$$(1.3') \quad \overline{\lim}_{n \rightarrow \infty} \{\sqrt{n} \cdot \text{dist}[\{z_{k,n}\}_{k=1}^n; d_{\infty}]\} \leq 2e = 5,43656 \dots$$

Azt azonban már nem vizsgálták az eddigi irodalomban, hogy az (1.3) alatti felső korlát a *lehető legjobb-e*, mint n függvénye midőn $n \rightarrow \infty$. Newman és Rivlin [8], valamint Saff és Varga [10] eredményei alapján könnyen megmutatható, hogy

4.3 ábra: $s_{27}(27z)$ gyökei.

az (1.3) alatti felső korlát, mint n függvénye, valóban a *lehető legjobb*. Pontosabban, Carpenter, Varga és Waldvogel [4] bebizonyították az alábbiakat (a bizonyítás megtalálható 4.2. alatt).

2. ÁLLÍTÁS ([4]). Ha $\{z_{k,n}\}_{k=1}^n$ jelöli az $s_n(nz)$ gyökeket, és t_1 jelöli (lásd (2.2)) az $\operatorname{erfc}(w) := \frac{2}{\sqrt{\pi}} \int_w^\infty e^{-t^2} dt$ (ahol $\operatorname{Im} t_1 > 0$) origóhoz legközelebbi gyökét, akkor az (1.3) alatti eredmény a *lehető legjobb*, mint n függvénye, amennyiben

$$(1.4) \quad \lim_{n \rightarrow \infty} \{\sqrt{n} \cdot \operatorname{dist}[\{z_{k,n}\}_{k=1}^n; D_\infty]\} \geq (\operatorname{Im} t_1 + \operatorname{Re} t_1) = 0,63665 \dots,$$

A 4.2 és 4.3 ábrákat megvizsgálva megállapíthatjuk, hogy nyilvánvalóan azon $\{z_{k,n}\}_{k=1}^n$ gyökök konvergenciája (a D_∞ görbéhez) a *gyorsabb*, amelyek egyenletesen távol maradnak a $z = 1$ ponttól. Valóban, ha a $z = 1$ pont körüli C_δ kört, vagyis a

$$(1.5) \quad C_\delta := \{z \in \mathbb{C} : |z - 1| < \delta\} \quad (0 < \delta \leq 1)$$

halmazt használjuk a $z = 1$ -hez közeli $\{z_{k,n}\}_{k=1}^n$ pontok *kizárására*, akkor ez a megfigyelt gyorsabb konvergencia *kvantitatív*á tehető. Pontosabban, a 4.2. bebizonyítjuk az alábbi eredményt.

3. TÉTEL ([4]). Ha $\{z_{k,n}\}_{k=1}^n$ jelöli a normalizált $s_n(nz)$ részletösszegek gyökeinek halmazát, és δ egy tetszőleges rögzített szám úgy, hogy $0 < \delta \leq 1$, akkor

$$(1.6) \quad \text{dist}[\{z_{k,n}\}_{k=1}^n \setminus C_\delta; D_\infty] = O\left(\frac{\log n}{n}\right) \quad (n \rightarrow \infty).$$

Továbbá az (1.6) alatti $O\left(\frac{\log n}{n}\right)$ korlát a lehető legjobb, mint n függvénye abban az értelemben, hogy

$$(1.6') \quad \lim_{n \rightarrow \infty} \left\{ \frac{n}{\log n} \cdot \text{dist}[\{z_{k,n}\}_{k=1}^n \setminus C_\delta; D_\infty] \right\} \leq 0,10900 \dots > 0$$

bármely $0 < \delta \leq 1$ esetén.

A 4.2 és 4.3 ábrákon az $s_n(nz)$ függvény $\{z_{k,n}\}_{k=1}^n$ gyökeinek szemmel látható „reguláris” eloszlására azt sejteti, hogy talán létezik a Δ egységkörlemezben egy (n -től függő) D_n sima ív, amely sokkal jobb approximációt szolgáltat az $s_n(nz)\{z_{k,n}\}_{k=1}^n$ gyökeire, mint a D_∞ görbe. Szegő [12] munkájával összhangban legyen

$$(1.7) \quad D_n := \left\{ z \in \mathbb{C} : |ze^{1-z}|^n = \tau_n \sqrt{2\pi n} \left| \frac{1-z}{z} \right|, |z| \leq 1, \right. \\ \left. \text{és } |\arg z| \geq \cos^{-1} \left(\frac{n-2}{n} \right) \right\}$$

bármely $n \geq 1$ esetén, ahol a Stirling formula (lásd Henrici [6, 377. oldal]) szerint

$$(1.8) \quad \tau_n := \frac{n!}{n^n e^{-n} \sqrt{2\pi n}} \approx 1 + \frac{1}{12n} + \frac{1}{288n^2} \\ - \frac{139}{51480n^3} + \dots \quad (n \rightarrow \infty).$$

Megemlítjük, hogy $\log \tau_n$ kifejezhető (lásd Henrici [6, 377. oldal]) a $J(z)$ Binet függvény segítségével, ami az alábbi aszimptotikus sorfejtéssel rendelkezik (lásd [6, 359. oldal]):

$$\log \tau_n = J(n) \approx \frac{n^{-1}}{12} - \frac{n^{-3}}{360} + \frac{n^{-5}}{1260} - \frac{n^{-7}}{1680} + \frac{n^{-9}}{1188} - \dots \quad (n \rightarrow \infty).$$

Továbbá megmutatható, hogy az (1.7) alatti D_n ív (lásd [4, Prop. 3]) egy jól definiált görbe, amely a Δ zárt egységkörlemez belsejében fekszik bármely pozitív egész n esetén. Úgyszintén érdemes megjegyezni, hogy (lásd 4, Prop. 3]) a D_n ív bármely n esetén csillagszerű az origóra vonatkozólag, vagyis bármely rögzített valós θ esetén, hacsak $|\theta| \geq \cos^{-1} \left(\frac{n-2}{n} \right)$, a $\{z = re^{i\theta} : r \geq 0\}$ sugár egyetlen pontban metszi D_n -et. Ugyancsak megemlítjük, hogy a D_n (1.7) alatti definíciójában az $|\arg z| \geq \cos^{-1} \left(\frac{n-2}{n} \right)$ korlátozás abból a tényből fakad, hogy $s_n(nz)$ összes gyöke az $|\arg z| \geq \cos^{-1} \left(\frac{n-2}{n} \right)$ szektorban van (lásd Saff és Varga [9]). Az (1.7) alatti D_n ívvel kapcsolatban megfogalmazzuk az alábbi új eredményt, amelyet a 4.3 pontban bizonyítunk be.

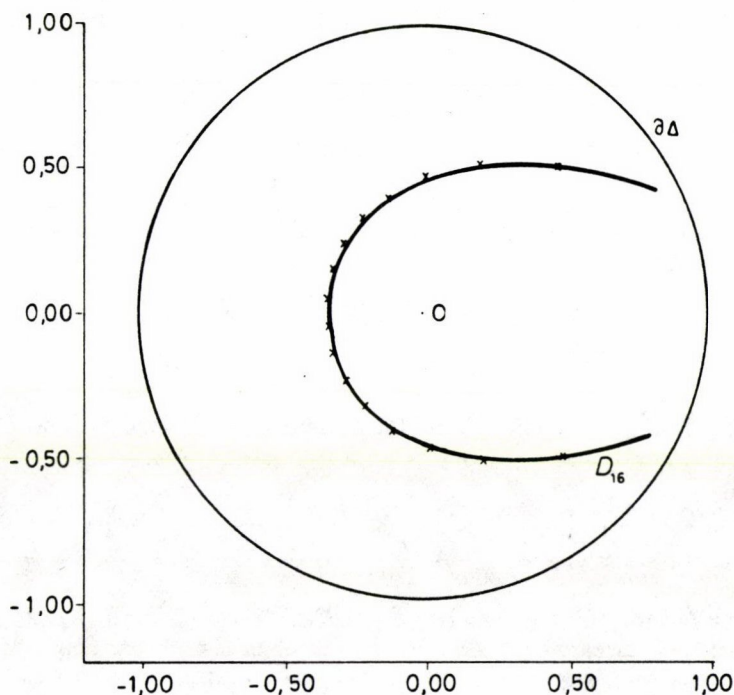
4. TÉTEL ([4]). Ha $\{z_{k,n}\}_{k=1}^n$ jelöli a normalizált $s_n(nz)$ részletösszegek gyökeit, és δ egy tetszőleges rögzített szám úgy, hogy $0 < \delta \leq 1$, akkor

$$(1.9) \quad \text{dist}[\{z_{k,n}\}_{k=1}^n \setminus C_\delta; D_n] = O\left(\frac{1}{n^2}\right) \quad (n \rightarrow \infty).$$

Továbbá az (1.9) alatti $O\left(\frac{1}{n^2}\right)$ korlát a lehető legjobb, mint az n függvénye, abban az értelemben, hogy

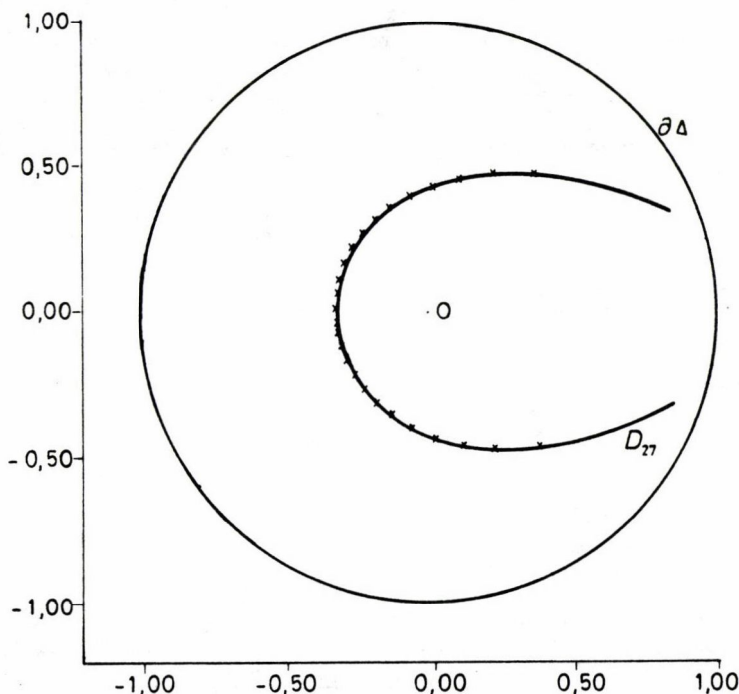
$$(1.9') \quad \lim_{n \rightarrow \infty} \{n^2 \cdot \text{dist}[\{z_{k,n}\}_{k=1}^n \setminus C_\delta; D_n]\} \geq 0,13326 \dots > 0$$

bármely $0 < \delta \leq 1$ esetén.



4.4 ábra: D_{16} és $s_{16}(16z)$ gyökei.

A 4. Tétel eredményének illusztrálására a 4.4 ábrán feltüntettük az $s_{16}(16z)$ 16 gyökét, a D_{16} görbével együtt. Ugyanezt tettük a 4.5 ábrán az $s_{27}(27z)$ 27 gyökével és a D_{27} görbével kapcsolatban. A rajz pontosságának határában belül úgy tűnik, hogy az $s_{16}(16z)$ és $s_{27}(27z)$ gyökei a D_{16} és D_{27} görbéken helyezkednek el!



4.5 ábra: D_{27} és $s_{27}(27z)$ gyökei.

4.2. A 2. Állítás és a 3. Tétel bizonyítása

A 2. Állítás Bizonyítása. Mint azt Newman és Rivlin [8] megmutatta,

$$(2.1) \quad \left\{ \frac{s_n(n + \sqrt{2n} w)}{e^{n + \sqrt{2n} w}} \right\}_{n=1}^{\infty} \text{ egyenletesen konvergál a } \frac{1}{\sqrt{\pi}} \int_w^{\infty} e^{-t^2} dt =: \frac{1}{2} \operatorname{erfc}(w) \text{ határértékhez}$$

midőn $n \rightarrow \infty$. A felső komplex számsík (vagyis $\operatorname{Im} w \geq 0$) bármely kompakt részhalmazán, midőn $n \rightarrow \infty$. Ha t_1 jelöli az $\operatorname{erfc}(w)$ origóhoz legközelebbi gyökét a felső félsíkon, akkor a numerikus analízisből ismert (lásd Fettis, Caslin és Cramer [5]), hogy

$$(2.2) \quad t_1 = -1,35481 \dots + i1,99146 \dots$$

Mivel az $s_n(n + \sqrt{2n} w)e^{-n - \sqrt{2n} w}$ gyökei kizárólag a w olyan értékei, amelyekre $s_n(n + \sqrt{2n} w)$ eltűnik, a (2.1) alatti egyenletes konvergencia Hurwitz tétele értelmében maga után vonja (lásd Titchmarsh [13, 119. oldal]), hogy $s_n(n + \sqrt{2n} w)$

rendelkezik egy gyökkel bármely kis, t_1 középpontú zárt körben, minden elég nagy n esetén. Következésképpen, mint azt Saff és Varga [10] megmutatta, $s_n(n + \sqrt{2n}w)$, rendelkezik egy

$$n + \sqrt{2n}(t_1 + o(1)) = n \left\{ 1 + \sqrt{\frac{2}{n}}(t_1 + o(1)) \right\} \quad (n \rightarrow \infty)$$

alakú gyökkel, vagy másszóval, $s_n(nz)$ -nek van egy

$$(2.3) \quad z_{1,n} := 1 + \sqrt{\frac{2}{n}}(t_1 + o(1)) \quad (n \rightarrow \infty)$$

alakú gyöke.

Mármost, ha z a D_∞ göbén van úgy, hogy $\operatorname{Re} z = 1 - \delta$, ahol $\delta > 0$ kicsi, akkor a D_∞ (1.1) alatti definíciójából közvetlenül következik, hogy $(1 - \delta)^2 + (\operatorname{Im} z)^2 = e^{-2\delta}$, vagyis

$$(2.4) \quad |\operatorname{Im} z| = \delta \left\{ 1 - \frac{2}{3}\delta + O(\delta^2) \right\} \quad (\delta \rightarrow 0).$$

Megjegyezzük, hogy (2.4)-ből következően a felső félsíkban a D_∞ görbe a valós tengellyel $\pi/4$ szöget lakot a $z = 1$ pontban. (Ez látható a 4.2. és 4.3. ábrákon is.) A (2.3) és (2.4) alapján számításokkal megmutatható (lásd Carpenter [3], 137. oldal), hogy nagy n esetén a (2.3) alatti $z_{1,n}$ távolsága a D_∞ görbétől kielégíti a

$$\operatorname{dist}[z_{1,n}; D_\infty] = \frac{1}{\sqrt{n}} \{\operatorname{Im} t_1 + \operatorname{Re} t_1 + o(1)\} \quad (n \rightarrow \infty)$$

feltételt, ahonnan

$$(2.5) \quad \lim_{n \rightarrow \infty} \{\sqrt{n} \operatorname{dist}[z_{1,n}; D_\infty]\} = \operatorname{Im} t_1 + \operatorname{Re} t_1 = 0,63665 \dots,$$

ahol az utóbbi eredményben felhasználtuk a (2.2) alatti numerikus becsléseket. Viszont (1.2) szerint, mivel $\operatorname{dist}[z_{1,n}; D_\infty] \leq \operatorname{dist}[\{z_{k,n}\}_{k=1}^n; D_\infty]$, következik a kívánt (1.4) alatti redmény. \square

A 4.1. pontbeli 3. Tétel bizonyításához a következő konstrukcióra van szükségünk. Az $s_n(z) := \sum_{j=1}^n z^j/j!$ definícióból differenciálással kapjuk, hogy

$$(2.6) \quad e^{-z} s_n(z) = 1 - \frac{1}{n!} \int_0^z \zeta^n e^{-\zeta} d\zeta.$$

Itt a ζ -t $n\zeta$ -val és a z -t nz -vel helyettesítve kapjuk, hogy

$$(2.7) \quad e^{-nz} s_n(nz) = 1 - \frac{n^{n+1}}{n!} \int_0^z \zeta^n e^{-n\zeta} d\zeta.$$

Innen a τ_n (1.8) alatti definícióját használva adódik, hogy

$$(2.8) \quad e^{-nz} s_n(nz) = 1 - \frac{\sqrt{n}}{\tau_n \sqrt{2\pi}} \int_0^z (\zeta e^{1-\zeta})^n d\zeta.$$

Mármost Szegő [12] alapján tudjuk, hogy $w = \zeta e^{1-\zeta}$ egyrétű a $|\zeta| < 1$ körben. (Ennek bizonyítására lásd a $\sigma = 0$ speciális esetet Saff és Varga [11, Lemma 4.1] munkájában.) Mivel végső soron minket csak az $s_n(nz)$ gyökei érdekelnek (amelyek 4.1. pont alapján a $|z| < 1$ körben fekszenek bármely $n > 1$ esetén), a (2.8)-ban elvégezzük a $w = \zeta e^{1-\zeta}$ változó-transzformációt, s azt kapjuk, hogy

$$(2.9) \quad e^{-nz} s_n(nz) = 1 - \frac{\sqrt{n}}{\tau_n \sqrt{2\pi}} \int_0^{ze^{1-z}} w^{n-1} \left(\frac{\zeta(w)}{1-\zeta(w)} \right) dw.$$

A fenti integrál alakja az alábbi megfontolást sugallja (megint csak Szegő [12] eredeti munkája alapján). Tekintsük az

$$(2.10) \quad \int_0^A w^{n-1} G(w) dw$$

integrált, ahol az integrációs út a 0 és A pontokat összekötő komplex egyenes szakasz. Feltéve, hogy $G(w)$ analitikus egy, ezen $[0, A]$ szakaszt tartalmazó nyílt tartományban, a $G(w)$ -t Taylor sorba fejtve a $w = A$ pont körül azt kapjuk, hogy

$$(2.11) \quad \int_0^A w^{n-1} G(w) dw = \sum_{m=0}^{\infty} \frac{G^{(m)}(A)}{m!} \int_0^A w^{n-1} (w-A)^m dw.$$

Mivel ezen összeg m -ik tagjában szereplő integrál (egy változó-transzformáció után) nem más, mint a béta-integrál (lásd [6, 55. oldal]), e tagra teljesül a

$$\frac{G^{(m)}(A)}{m!} \int_0^A w^{n-1} (w-A)^m dw = \frac{(-1)^m A^{m+n} G^{(m)}(A)}{\prod_{j=0}^m (n+j)}$$

reláció bármely $m \geq 0$ és $n \geq 1$ esetén. Így a (2.10) alatti integrál az

$$(2.12) \quad \int_0^A w^{n-1} G(w) dw = A^n \sum_{m=0}^{\infty} \frac{(-1)^m A^m G^{(m)}(A)}{\prod_{j=0}^m (n+j)}$$

alakban írható.

A (2.9) és (2.10) alatti integrálok összekapcsolhatók az $A := ze^{1-z}$, $F(\zeta) := \zeta/(1-\zeta)$ és $G(w) := F(\zeta(w))$ jelöléseket használva, ahol $w := \zeta e^{1-\zeta}$. Ha z bármely

belső pontja a Δ zárt egységgörnek, akkor az így definiált $G(w)$ valóban analitikus egy, a $[0, ze^{1-z}]$ szakaszt tartalmazó nyílt tartományban, és fennáll a (2.12) reprezentáció. A $G^{(m)}(ze^{1-z})$ explicit értékeit kiszámítva $0 \leq m \leq 4$ esetén, megadhatjuk a (2.12) első néhány tagját

$$(2.13) \quad \int_0^{ze^{1-z}} w^{n-1} \left(\frac{\zeta(w)}{1-\zeta(w)} \right) dw = \frac{z(ze^{1-z})^n}{n(1-z)} \times \left\{ 1 - \frac{1}{(n+1)(1-z)^2} + \frac{z(4-z)}{(n+1)(n+2)(1-z)^4} - \frac{z^2(27-14z+2z^2)}{(n+1)(n+2)(n+3)(1-z)^6} + \frac{z^3(256-203z+58z^2-6z^3)}{(\prod_{j=1}^4 (n+j))(1-z)^8} - \dots \right\} \quad (n=1, 2, \dots).$$

A (2.11) alatti Taylor sorfejtés szeleteinek Cauchy-féle maradéktagját megbecsülve, (2.12) alapján láthatjuk, hogy bármely nemnegatív egész N esetén

$$(2.14) \quad \int_0^{ze^{1-z}} w^{n-1} \left(\frac{\zeta(w)}{1-\zeta(w)} \right) dw = (ze^{1-z})^n \sum_{m=0}^N \frac{(-1)^m (ze^{1-z})^m G^{(m)}(ze^{1-z})}{\prod_{j=0}^m (n+j)} + O\left(\frac{1}{n^{N+2}}\right) \quad (n \rightarrow \infty)$$

egyenletesen a $\Delta \setminus \{1\}$ bármely kompakt részhalmazán. (Ezen eredményt természetesen Szegő [12] munkája motiválja, ahol (2.14) az $N=0$ esetben van bebizonyítva.)

A (2.14) és (2.13) formulákat $N=0$ esetén alkalmazva kapjuk, hogy

$$(2.15) \quad \int_0^{ze^{1-z}} w^{n-1} \left(\frac{\zeta(w)}{1-\zeta(w)} \right) dw = \frac{z(ze^{1-z})^n}{n(1-z)} \left\{ 1 + O\left(\frac{1}{n}\right) \right\} \quad (n \rightarrow \infty)$$

egyenletesen a $\Delta \setminus \{1\}$ bármely kompakt részhalmazán. Így, ha z egy gyöke az e^z normalizált $s_n(nz)$ részletösszegének, akkor (2.9) és (2.15) alapján adódik, hogy

$$1 - \frac{\sqrt{n}}{\tau_n \sqrt{2\pi}} \cdot \frac{z(ze^{1-z})^n}{n(1-z)} \left\{ 1 + O\left(\frac{1}{n}\right) \right\} = 0,$$

vagy ezzel ekvivalens módon

$$(2.16) \quad (ze^{1-z})^n = \tau_n \sqrt{2\pi n} \left(\frac{1-z}{z} \right) \left\{ 1 + O\left(\frac{1}{n}\right) \right\} \quad (n \rightarrow \infty)$$

egyenletesen a $\Delta \setminus \{1\}$ bármely kompakt részhalmazán.

Ezzel elérkeztünk a 3. Tétel bizonyításához.

A 3. Tétel bizonyítása. Szegő [12] munkájából ismert, hogy a $w = ze^{1-z}$ függvény a D_∞ Szegő-féle görbét konformisan leképezi a $|w| < 1$ kör belsejére, továbbá a D_∞ pontjait egy-egyértelmű módon képezi le a $|w| = 1$ körvonalra úgy, hogy a $w = ze^{1-z}$ argumentuma monoton növekszik 0-tól 2π -ig, midőn z a $z = 1$ pontból kiindulva végighalad a D_∞ görbén pozitív értelemben. Szegő [12] úgyszintén megmutatta, hogy az $s_n(nz)$ zérushelyei aszimptotikusan *egyenletesen oszlanak el* szögtartományok szerint a $w = ze^{1-z}$ leképezés mellett (midőn $n \rightarrow \infty$). Pontosabban, legyenek π_1 és ϕ_2 tetszőleges valós számok úgy, hogy $0 < \phi_1 < \phi_2 < 2\pi$, és legyenek z_1 és z_2 rendre a $w_1 = e^{i\phi_1}$ és $w_2 = e^{i\phi_2}$ invers képei a $w = ze^{1-z}$ leképezésre vonatkozólag, úgyhogy z_1 és z_2 a D_∞ görbén vannak, s emellett $0 < \arg z_1 < \arg z_2 < 2\pi$. A z -síkbeli S szögtartományt definiáljuk a

$$S = \{z \in \mathbb{C} : \arg z_1 \leq \arg z \leq \arg z_2\}$$

relációval. Ekkor, ha σ_n jelöli az $s_n(nz)$ gyökeinek számát S -ben, akkor Szegő [12] megmutatta, hogy

$$\lim_{n \rightarrow \infty} \frac{\sigma_n}{n} = \frac{\phi_2 - \phi_1}{2\pi}.$$

Ebből következik, hogy nagy n esetén az $s_n(nz)$ gyökei szög szerint durván *egyenletesen oszlanak el* a w -síkbán, a $w = ze^{1-z}$ leképezésre nézve.

Ezt a következőképpen használhatjuk fel. Ha tekintjük az n számú, a $|w| = 1$ körön egyenletes eloszlású $\{\exp[i(2k-1)\pi/n]\}_{k=1}^n$ pontot, akkor legyenek $\{\tilde{z}_{k,n}\}_{k=1}^n$ ezek egyértelmű inverzei a z -síkon a $w = ze^{1-z}$ leképezésre vonatkozólag, vagyis

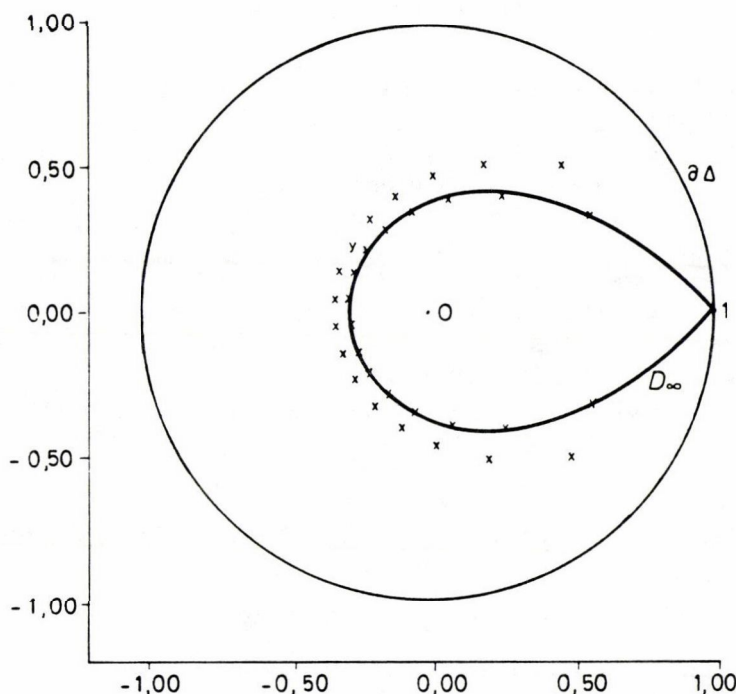
$$(2.17) \quad \tilde{z}_{k,n} e^{1-\tilde{z}_{k,n}} = \exp[i(2k-1)\pi/n] \quad (k = 1, 2, \dots, n).$$

Definíció szerint a $\{\tilde{z}_{k,n}\}_{k=1}^n$ pontok a D_∞ görbén fekszenek; a 4.6. ábrán, mint *-ok szemléltettük a $\{\tilde{z}_{k,16}\}_{k=1}^{16}$ pontokat a D_∞ görbén az $s_{16}(16z)$ függvény $\{z_{k,16}\}_{k=1}^{16}$ gyökeivel együtt. Ugyanezt tettük a 4.7. ábrán a $\{\tilde{z}_{k,27}\}_{k=1}^{27}$ értékekkel valamint az $s_{27}(27z)$ függvény $\{z_{k,27}\}_{k=1}^{27}$ gyökeivel kapcsolatban. A (2.17)-ből következőleg

$$(2.18) \quad (\tilde{z}_{k,n} e^{1-\tilde{z}_{k,n}})^n = -1 \quad (k = 1, 2, \dots, n).$$

Ami a $z_{k,n}$ $\tilde{z}_{k,n}$ approximációját illeti, legyen $z_{k,n}$, az $s_n(nz)$ egy gyöke $z_{k,n} = \tilde{z}_{k,n} + \delta_{k,n}$ alakú, és helyettesítjük ezt (2.16)-ba. A (2.18) felhasználásával, egyszerű számítással logaritmusokat véve és n -nel osztva kapjuk, hogy

$$(2.19) \quad \begin{aligned} & - \left(1 - \frac{1}{\tilde{z}_{k,n}}\right) \delta_{k,n} + O(\delta_{k,n}^2) = \\ & = \frac{1}{n} \log \left\{ \tau_n \sqrt{2\pi n} \left(1 - \frac{1}{\tilde{z}_{k,n}}\right) \right\} + O\left(\frac{\delta_{k,n}}{n}\right) + O\left(\frac{1}{n^2}\right), \end{aligned}$$



4.6 ábra: $\{z_{k,16}\}_{k=1}^{16}$ és $\{\tilde{z}_{k,16}\}_{k=1}^{16}$.

feltéve, hogy csak azokat a $z_{k,n}$ gyököket tekintjük, amelyek a C_δ -n kívül fekszenek (ahol δ egy rögzített szám úgy, hogy $0 < \delta \leq 1$, és C_δ az (1.5) által definiált). (A (2.19)-beli logaritmus esetén azt az egyértékű kiterjesztést vesszük a $\mathbb{C} \setminus [0, +\infty)$ halmazon, amelyre $\log(-1) = i\pi$.)

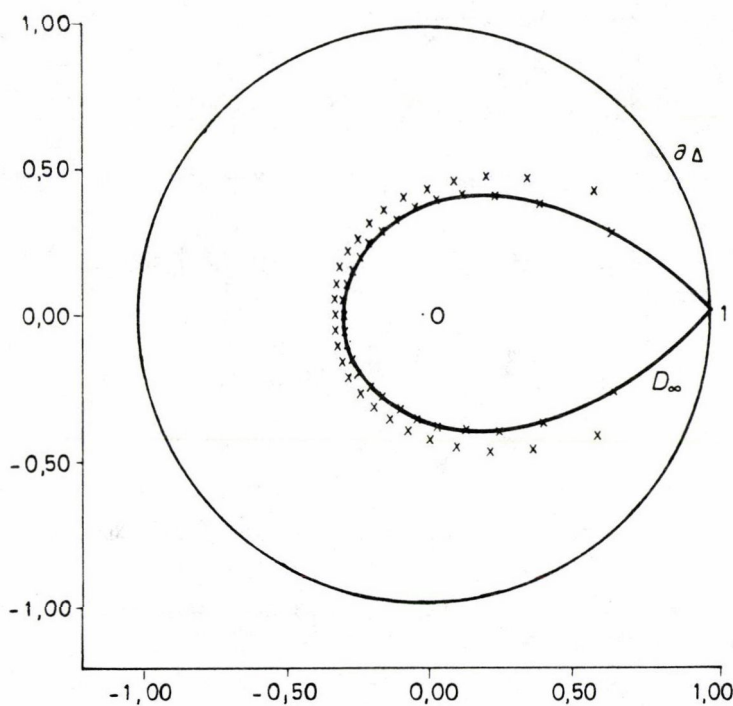
A $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ halmazhoz nyilvánvalóan létezik egy pozitív, csak a δ -tól függő c_1 konstans úgy, hogy

$$0 < c_1 \leq \left| 1 - \frac{1}{\tilde{z}_{k,n}} \right|$$

bármely, a $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ pontokhoz rendelt $\tilde{z}_{k,n}$ pontokra. Így (2.19)-ből következik, hogy

$$(2.20) \quad \delta_{k,n} = O\left(\frac{\log[\tau_n \sqrt{2\pi n}]}{n}\right) = O\left(\frac{\log n}{n}\right) \quad (n \rightarrow \infty)$$

bármely $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ -beli pontra. De mivel $\tilde{z}_{k,n}$ nem szükségképpen a legközelebbi pont a $z_{k,n}$ -hez a D_∞ görbén, ezért $\text{dist}[z_{k,n}; D_\infty] \leq |\delta_{k,n}|$ a $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ valamennyi pontjára, amiből (2.20) alapján következik a 3. Tétel kívánt (1.6) relációja.


 4.7 ábra: $\{z_{k,27}\}_{k=1}^{27}$ és $\{\tilde{z}_{k,27}\}_{k=1}^{(27z)}$.

Annak érdekében, hogy megmutassuk, hogy a 3. Tételben az (1.6) nagyságrend *a lehető legjobb*, mint az n függvénye, először legyen $n = 2m + 1$ egy tetszőleges páratlan szám, és legyen $-\mu$ a D_∞ Szegő-görbe *negatív* valós pontja, vagyis legyen μ a $\mu e^{1+\mu} = 1$ egyenlet egyértelmű pozitív gyöke; numerikusan $\mu = 0,27846 \dots$. A (2.18)-ból láthatjuk, hogy

$$(2.21) \quad \tilde{z}_{m+1,2m+1} = -\mu = -0,27846 \dots \quad (m = 0, 1, \dots),$$

és hasonlóképpen adódik, hogy $z_{m+1,2m+1}$ a $s_{2m+1}((2m+1)z)$ páratlan polinom egyértelmű (negatív) gyöke. A (2.21) felhasználásával, (2.19)-ből következik, hogy

$$(2.22) \quad \lim_{m \rightarrow \infty} \left\{ \left(\frac{2m+1}{\log(2m+1)} \right) \cdot \delta_{m+1,2m+1} \right\} = -\frac{1}{2(1+\frac{1}{\mu})} = -0,10890 \dots$$

Mivel $z_{m+1,2m+1}$ a D_∞ görbén *kívül* esik, ezért $\delta_{m+1,2m+1} < 0$ és $-\delta_{m+1,2m+1} = \text{dist}[z_{m+1,2m+1}; D_\infty]$ továbbá $z_{m+1,2m+1}$ a C_δ görbén *kívül* fekszik *bármely* olyan

δ -ra, amelyre $0 < \delta \leq 1$ teljesül, Így (2.22) a

$$(2.23) \quad \lim_{m \rightarrow \infty} \left\{ \left(\frac{2m+1}{\log(2m+1)} \right) \cdot \text{dist}[z_{m+1,2,+1}; D_\infty] \right\} = \\ = \frac{1}{2(1 + \frac{1}{\mu})} = 0,10890 \dots$$

alakot ölti. Másrészt, mivel $\text{dist}[z_{m+1,2m+1}; D_\infty] \leq \text{dist}[\{z_{k,2m+1}\}_{k=1}^{2m+1} \setminus C_\delta; D_\infty]$, ezért (2.23) szerint

$$(2.24) \quad \lim_{m \rightarrow \infty} \left\{ \left(\frac{2m+1}{\log(2m+1)} \right) \cdot \text{dist}[\{z_{k,2m+1}\}_{k=1}^{2m+1} \setminus C_\delta; D_\infty] \right\} \geq \\ \geq \frac{1}{2(1 + \frac{1}{\mu})} = 0,10890 \dots > 0.$$

Az $n = 2m$ esetben hasonlóképpen megmutatható, hogy a (2.23) analogonja

$$(2.25) \quad \lim_{m \rightarrow \infty} \left\{ \frac{2m}{\log(2m)} \cdot \text{dist}[z_{m,2m}; D_\infty] \right\} = \frac{1}{2(1 + \frac{1}{\mu})} = 0,10890 \dots,$$

úgyhogy (lásd (2.24))

$$(2.26) \quad \lim_{m \rightarrow \infty} \left\{ \frac{2m}{\log(2m)} \cdot \text{dist}[\{z_{k,2m}\}_{k=1}^{2m} \setminus C_\delta; D_\infty] \right\} \geq \frac{1}{2(1 + \frac{1}{\mu})} > 0.$$

A (2.24) és (2.25) kombinációja azt adja, hogy

$$(2.27) \quad \lim_{m \rightarrow \infty} \left\{ \frac{n}{\log n} \cdot \text{dist}[\{z_{k,n}\}_{k=1}^n \setminus C_\delta; D_\infty] \right\} \geq 0,10890 \dots > 0,$$

ami nem más, mint a 3. Tétel kívánt (1.6') relációja. \square

4.3. A 4. Tétel bizonyítása

A (2.16) alatti reláció egyenletesen teljesül n -ben, midőn $n \rightarrow \infty$, az $s_n(nz)$ bármely olyan z gyökére, amely a $\Delta \setminus \{1\}$ valamely kompakt részhalmazában fekszik. A (2.16)-ban abszolút értékeket és n -dik gyököket véve teljesen világos, hogy $n \rightarrow \infty$ esetén az (1.1) alatti Szegő-görbe *miért* jelenik meg úgy, mint az *egyetlen* lehetséges hely, ahol az $\{s_n(nz)\}_{n=1}^\infty$ gyökei aszimptotikusan torlódhatnak. Mint azt a 2. Állításból tudjuk, az $s_n(nz)$ gyökeinek maximális távolsága a D_∞ görbétől $O\left(\frac{1}{\sqrt{n}}\right)$ midőn $n \rightarrow \infty$, és ez a maximális távolság a 3. Tételben az $O\left(\frac{\log n}{n}\right)$ értékre javítható a $\Delta \setminus \{1\}$ kompakt részhalmazain.

Másrészről természetes dolog azt kérdezni, hogy lehetséges-e definiálni egy ívet, mondjuk D_n -et n -től függően úgy, hogy az $s_n(nz)$ gyökei *lényegesen* közelebb fesszenek D_n -hez, mint a D_∞ görbéhez. Természetesen bármely, az $s_n(nz)$ $\{z_{k,n}\}_{k=1}^n$ gyökein átvezető sima görbe megfelelne. Kiderült, hogy egy ilyen D_n ívet lehet *definálni* anélkül, hogy az $s_n(nz)$ $\{z_{k,n}\}_{k=1}^n$ gyökeit explicite ismernénk. Valóban, az (1.7) alatti D_n görbe definíciójához úgy is eljuthatunk, hogy (2.16)-ban elhagyjuk az $O\left(\frac{1}{n}\right)$ tagot, és mindenütt abszolút értéket veszünk!

Ezzel elérkeztünk a 4. Tétel bizonyításához.

A 4. Tétel bizonyítása. Egy tetszőleges rögzített olyan δ mellett, amelyre $0 < \delta \leq 1$, tekintsük a $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ halmazt, ahol C_δ az (1.5) által definiált, és ahol ismét csak $\{z_{k,n}\}_{k=1}^n$ jelöli az $s_n(nz)$ gyökeinek halmazát. Ekkor a $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ halmazbeli bármely $z_{k,n}$ gyökre teljesül (2.6), vagyis

$$(3.1) \quad (ze^{1-z})^n = \tau_n \sqrt{2\pi n} \left(\frac{1-z}{z} \right) \left\{ 1 + O\left(\frac{1}{n}\right) \right\} \quad (n \rightarrow \infty),$$

ahol az $O\left(\frac{1}{n}\right)$ tagban rejlő konstans csak a δ -tól függ. Másrészt a D_n ívre vonatkozó megjegyzésünk értelmében bármely olyan θ esetén, melyre $\theta_n \leq \theta \leq 2\pi - \theta_n$ teljesül, ahol $\theta_n := \cos^{-1}\left(\frac{n-2}{n}\right)$ tetszőleges $n \geq 1$ mellett, létezik egy egyértelmű $r_n(\theta)$ a $(0,1)$ intervallumban úgy, hogy $z = r_n(\theta)e^{i\theta}$ a D_n görbén felszik. Ez (1.7) alapján azt jelenti, hogy létezik egy $\Psi(n, \theta)$ valós szám úgy, hogy

$$(3.2) \quad \frac{z(ze^{1-z})^n}{\tau_n \sqrt{2\pi n}(1-z)} = e^{i\Psi(n, \theta)},$$

ahol

$$(3.3) \quad \Psi(n, \theta) := n[\theta - r_n(\theta) \sin \theta] + \theta + \tan^{-1} \left[\frac{r_n(\theta) \sin \theta}{1 - r_n(\theta) \cos \theta} \right],$$

bármely $\theta_n \leq \theta \leq 2\pi - \theta_n$ esetén. Megmutatható, hogy ez a $\Psi(n, \theta)$ egy szigorúan monoton növekvő függvénye a θ -nak a $[\theta_n, 2\pi - \theta_n]$ intervallumon, és hogy létezik pontosan n különböző θ érték a $(\theta_n, 2\pi - \theta_n)$ intervallumban, amelyre $\Psi(n, \theta) \equiv 0 \pmod{2\pi}$. Ha a D_n ezen n pontját $\{\hat{z}_{k,n}\}_{k=1}^n$ jelöli, akkor (3.2)-ből következik, hogy

$$(3.4) \quad (\hat{z}_{k,n} e^{1-\hat{z}_{k,n}})^n = \tau_n \sqrt{2\pi n} \left(\frac{1-\hat{z}_{k,n}}{\hat{z}_{k,n}} \right) \quad (k = 1, 2, \dots, n).$$

Mivel $\{z_{k,n}\}_{k=1}^n$ az $s_n(nz)$ gyökeinek halmaza (növekvő argumentumokkal), $z_{k,n}$ -et kifejezhetjük a $z_{k,n} = \hat{z}_{k,n} + \delta_{k,n}$ alakban. Így a $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ bármely $z_{k,n}$ elemére (3.1) alapján teljesül, hogy

$$(3.5) \quad (z_{k,n} e^{1-z_{k,n}})^n = \tau_n \sqrt{2\pi n} \left(\frac{1-z_{k,n}}{z_{k,n}} \right) \left\{ 1 + O\left(\frac{1}{n}\right) \right\}.$$

A $z_{k,n}$ -et $\hat{z}_{k,n} + \delta_{k,n}$ -nel helyettesítve (3.5)-ben, és felhasználva (3.4)-et, logaritmusokat véve és n -nel osztva ez a következőre redukálódik:

$$(3.6) \quad \log \left(1 + \frac{\delta_{k,n}}{\hat{z}_{k,n}} \right) - \delta_{k,n} = \frac{1}{n} \log \left(1 - \frac{\delta_{k,n}}{1 - \hat{z}_{k,n}} \right) - \frac{1}{n} \log \left(1 + \frac{\delta_{k,n}}{\hat{z}_{k,n}} \right) + O \left(\frac{1}{n^2} \right),$$

midőn $n \rightarrow \infty$. Az itt fellépő különböző tagokat kifejtve (azzal a feltevéssel, hogy $\delta_{k,n}$ elég kicsi), könnyen adódik, hogy

$$(3.7) \quad \delta_{k,n} = O \left(\frac{1}{n^2} \right) \quad (n \rightarrow \infty)$$

a $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ bármely pontjára. Mivel $\hat{z}_{k,n}$ ismét csak szükségszerűen a D_n legközelebbi pontja $z_{k,n}$ -hez, $\text{dist}[z_{k,n}; D_n] \leq \delta_{k,n}$ bármely $\{z_{k,n}\}_{k=1}^n \setminus C_\delta$ halmazbeli pontra. Innen a 4. Tétel kívánt (1.9) relációja következik.

Végül annak megmutatására, hogy a (3.7) alatti eredmény a *lehető legjobb*, mint az n függvénye, ismét csak elegendő (mint a 4.2. pontban) az $s_n(nz)$ speciális $\{z_{m+1,2m+1}\}_{m=1}^\infty$ és $\{z_{m,2m}\}_{m=1}^\infty$ gyöksorozatait tekinteni. Emlékezve a (2.21) alatti μ számra, a 3. Tétel bizonyításának mintájára megmutatható, hogy

$$(3.8) \quad \lim_{n \rightarrow \infty} \{n^2 \cdot \text{dist}[\{z_{k,n}\}_{k=1}^n \setminus C_\delta; D_n]\} \geq \frac{\mu}{(1+\mu)^3} = 0,13326 \dots,$$

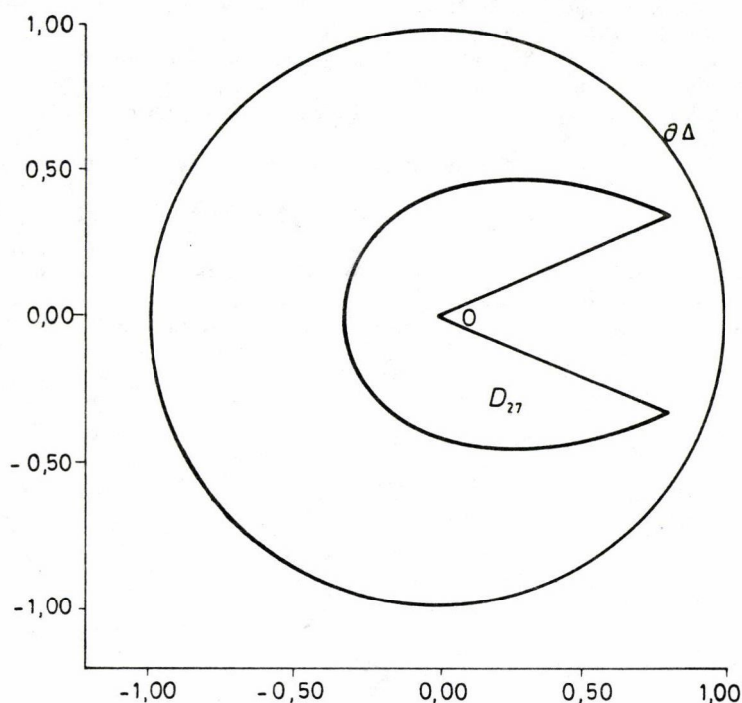
bármely olyan δ , amelyre $0 < \delta \leq 1$, ami a kívánt, (1.9')-ben megfogalmazott élességet bizonyítja. \square

E fejezet lezárásaként tekintsük a

$$(3.9) \quad \mathcal{D}_n := \left\{ z \in \mathbb{C} : |ze^{1-z}|^n \leq \tau_n \sqrt{2\pi n} |1-z|/|z|, \quad |z| \leq 1 \text{ és } |\arg z| \cos^{-1}((n-z)/n) \right\} \quad (n = 1, 2, \dots),$$

zár halmazt, amelynek határa az (1.7) alatti D_n ívből, valamint két, az origóból induló szakaszból áll. A 4.8. ábrán a \mathcal{D}_{27} halmazt szemléltettük. Megjegyezzük, hogy a \mathcal{D}_n halmaz nyilván csillagszerű az origóból nézve. Mivel tudjuk, hogy az $s_n(nz)$ $\{z_{k,n}\}_{k=1}^n$ gyökei a D_∞ Szegő-görbén kívül fekszenek, természetes dolog azt kérdezni, hogy vajon hasonlóképpen, e gyökök a D_n -en is kívül esnek-e bármely $n \geq 1$ esetén. Sajnos, a 4.4 és 4.5 ábrák nem elég pontosak ahhoz, hogy a választ megsejthessük. Nagypontosságú számításokra (120 értékes jegy), valamint aszimptotikus vizsgálatokra alapozva viszont kimondhatjuk az alábbi két eredményt:

- i) létezik egy legkisebb pozitív egész n_0 úgy, hogy
- $$(3.10) \quad \{z_{k,n}\}_{k=1}^n \cap \mathcal{D}_n \neq \emptyset \text{ bármely pozitív } n > n_0 \text{ esetén, és}$$
- ii) $\{z_{k,n}\}_{k=1}^n \cap \mathcal{D}_n = \emptyset$, bármely $n = 1, 2, \dots, 96$ mellett.


4.8 ábra: D_{27} .

Ezen túlmenően úgy tűnik, hogy $n_0 = 96$. Ezek az eredmények megtalálhatók [14]-ben.

IRODALOM

- [1] N. ANDERSON, E. B. SAFF és R. S. VARGA, „On the Eneström-Kakeya Theorem and its sharpness”, *Linear Algebra Appl.* **28** (1979), 5–16.
- [2] J. D. BUCKHOLTZ, „A characterization of the exponential series”, *Amer. Math. Monthly* **73** Part II (1966), 121–123.
- [3] A. J. CARPENTER, „Some theoretical and computational aspects of approximation theory”, *Ph. D. disszertáció* (The University of Leeds, Leeds, England, 1988).
- [4] A. J. CARPENTER, R. S. VARGA és J. WALDVOGEL, „Asymptotics for the zeros of the partial sums of e^z ”, *Rocky Mountain J. Math.* **21** (1991), 99–120.
- [5] H. E. FETTIS, J. C. CASLIN and K. R. CRAMER, „Complex zeros of the error function and of the complementary error function”, *Math. Comp.* **27** (1973), 401–404.
- [6] P. HENRICI, *Applied and Computational Complex Analysis*, vol. 2 kötet (John Wiley & Sons, New York, 1977).
- [7] M. MARDEN, „Geometry of Polynomials”, *Mathematical Surveys No. 3* (Amer. Math. Soc. Providence, R.I., 1966).
- [8] D. J. NEWMAN és T. J. RIVLIN, „The zeros of the partial sums of the exponential function”, *J. Approx. Theory* **5** (1972), 405–412.

- [9] E.B. SAFF és R. S. VARGA, „On the zeros and poles of Padé approximants to e^z ”, *Numer. Math.* **25** (1975), 1–14.
- [10] E. B. SAFF és R. S. VARGA, „Zero-free parabolic regions for sequences of polynomials”, *SIAM J. Math. Anal.* **7** (1976), 344–357.
- [11] E. B. SAFF és R. S. VARGA, „On the zeros and poles of Padé approximants to e^z III”, *Numer. Math.* **30** (1978), 241–266.
- [12] G. SZEGŐ, „Über eine Eigenschaft der Exponentialreihe”, *Sitzungsber. Berl. Math. Ges.* **23** (1924), 50–64.
- [13] E. C. TITCHMARSH, *The Theory of Functions*, 2. kiadás (Oxford University Press, London, 1950).
- [14] R.S. VARGA és A.J. CARPENTER, „Asymptotics for the zeros of the partial sums of e^z . II”, *Computational Methods and Function Theory*, St. Ruscheweyh, E. B. Saff, L. Salinas, R.S. Varga szerkesztő, *Lecture Notes in Mathematics* 1435 (Springer-Verlag, Heidelberg, 1990), 201–207.

5. FEJEZET

Valós és komplex legjobb racionális approximáció

5.1. A $\gamma_{m,n}$ számok

Ebben a fejezetben egy olyan, az utóbbi időkben érdeklődést kiváltó jelenséggel foglalkozunk, amely akkor lép fel, ha *komplex* racionális függvényeket szembeállítunk (azonos rendű) *valós* racionális függvényekkel *valós* folytonos függvények (egyenletes normában való) approximációjakor a *valós* $[-1, +1]$ intervallumon.

Legyen π_m^r ill. π_m^c a z vagy x változójú, legfeljebb m -edfokú valós, ill. komplex együtthatójú polinomok halmaza. A nemnegatív számok egy (m, n) párja esetén legyen $\pi_{m,n}^r$ ill. $\pi_{m,n}^c$ az olyan p/q alakú racionális függvények halmaza, amelyre $p \in \pi_m^r(\pi_m^c)$ ill. $q \in \pi_n^r(\pi_n^c)$. Az

$$I := [-1, +1],$$

jelöléssel élve, legyen $C_r(I)$ az összes folytonos, valós értékű függvények halmaza az I -n. Továbbá bármely $f \in C_r(I)$ esetén legyen

$$(1.1) \quad E_{m,n}^r(f) := \inf_{g \in \pi_{m,n}^r} \|f - g\|_{L_\infty(I)}; \quad E_{m,n}^c(f) := \inf_{g \in \pi_{m,n}^c} \|f - g\|_{L_\infty(I)},$$

ahol valamely valós vagy komplex értékű, az I intervallumon definiált h függvényre

$$\|h\|_{L_\infty(I)} := \sup\{|h(x)| : x \in I\}.$$

A jelenség, amit e fejezetben vizsgálni szándékozunk, a következő. Azt állítjuk, hogy a nemnegatív egészek bármely olyan (m, n) párja esetén, ahol $n \geq 1$, létezik egy $f \in C_r(I)$ úgy, hogy

$$(1.2) \quad \begin{cases} (i) & E_{m,n}^c(f) < E_{m,n}^r(f), \text{ és} \\ (ii) & \text{az } f \text{ legjobb egyenletes approximációja a } \pi_{m,n}^c \text{ halmazból az } I \\ & \text{intervallumon nem egyértelmű.} \end{cases}$$

Mivel a nemnegatív egészek bármely (m, n) párja esetén $\pi_{m,n}^r$ egy valódi részhalmaza $\pi_{m,n}^c$ -nek, (1.1) alapján világos, hogy $E_{m,n}^c(f) \leq E_{m,n}^r(f)$ bármely $f \in C_r(I)$ esetén. Mivel egy tetszőleges komplex szám két valós paraméterből áll, ezért bármely $r_{m,n} \in \pi_{m,n}^r$ elemhez hozzárendelhető egy $R_{m,n} \in \pi_{m,n}^c$, amely kétszer annyi valós paraméterből áll. Önmagában heurisztikusan már ez is azt sugallhatja, hogy (1.2i)-ben $E_{m,n}^c(f)$ durván legfeljebb $E_{m,n}^r(f)/2$ bizonyos $f \in C_r(I)$ esetén. Később látni fogjuk (1.19)-ben, hogy az $n = 0$ esettől eltekintve, amikor $E_{m,0}^c(f) = E_{m,0}^r(f)$ bármely $f \in C_r(I)$ és $m \geq 0$ esetén, ez a heurisztikusan kikövetkeztetett egyenlőtlenség (és ennél több is) légyegében *helyes*!

Ismeretes (lásd Meinardus [5, 161. oldal]), hogy bármely $f \in C_r(I)$ egy *egyértelmű* $r_{m,n} \in \pi_{m,n}^r$ legjobb racionális approximációval rendelkezik az I -n. Másrészt

Walsh [14, 356. oldal] adott egy példát egy folytonos, komplex értékű függvényre, nevezetesen $f(z) := z + z^{-1}$, amely a komplex sík egy bizonyos félhold-alakú halmazon *nem* rendelkezik egyértelmű egyenletes racionális approximációval a $\pi_{1,1}^c$ halmazon. Az, hogy ez a nem-egyértelműség bekövetkezhet *valós* intervallumon vett *valós* függvényekre is, mint (1.2ii)-ben, talán meglepi az olvasót.

Abból a célból, hogy az (1.2) mindkét részét szemléltető példát adjunk, először is felidézzük (lásd [5, 161. oldal]), hogy bármely $f \in C_r(I)$ és a nemnegatív egészek bármely (m, n) párjára az

$$(1.3) \quad E_{m,n}^r(f) = \|f - r_{m,n}\|_{L_\infty(I)}$$

feltételt kielégítő, egyértelműen létező $r_{m,n} = p/q \in \pi_{m,n}^r$ (ahol p és q -ról feltételezzük, hogy nem rendelkeznek közös faktorokkal) pontosan jellemezhető egy $\{\xi_j\}_{j=1}^\ell$ *alternációs halmazzal*, ahol $-1 \leq \xi_1 < \xi_2 < \dots < \xi_\ell \leq 1$, és (egy rögzített $\lambda = 1$ vagy $\lambda = -1$ értékkel)

$$(1.4) \quad f(\xi_j) - r_{m,n}(\xi_j) = \lambda(-1)^j E_{m,n}^r(f) \quad (j = 1, 2, \dots, \ell),$$

és amelyre

$$(1.5) \quad \ell \geq 2 + \max\{m + \deg q; n + \deg p\}.$$

(Itt megállapodunk abban, hogy ha $p \equiv 0$, akkor $\deg p := -\infty$ és $\deg q := 0$, úgyhogy (1.5)-ben ez esetben $\ell \geq 2 + m$. Az ℓ -et a $\{\xi_j\}_{j=1}^\ell$ alternációs halmaz *hosszának* is nevezzük.) Tekintsük most az $f(x) := x^2 \in C_r(I)$ speciális függvényt. Ha $r_{1,1}(x) := p(x)/q(x) \equiv 1/2$ bármely *valós* x esetén, akkor

$$(1.6) \quad E_{1,1}^r(x^2) = 1/2,$$

mivel az összes (1.4) és (1.5) alatti feltétel teljesül (az $\ell = 3$, $\lambda = -1$, $m = n = 1$, $\deg p = \deg q = 0$, $\xi_1 := -1$, $\xi_2 = 0$, és $\xi_3 := 1$ értékekkel). Másrészt tekintsük a

$$(1.7) \quad g_{1,1}(x) := \frac{x + (\sqrt{2} - 1)i}{x + i}$$

függvényt, amely a $\pi_{1,1}^c$ halmazhoz tartozik. Rövid számolás mutatja, hogy

$$(1.8) \quad \|x^2 - g_{1,1}(x)\|_{L_\infty(I)} = \sqrt{2} - 1 = 0,41421 \dots,$$

s így az (1.1) jelöléssel élve, $E_{1,1}^c(x^2) \leq 0,41421 \dots$. Ebből és (1.6)-ból következik, hogy

$$(1.9) \quad E_{1,1}^c(x^2) < E_{1,1}^r(x^2).$$

Mármost a nemnegatív egészek bármely (m, n) párjára és bármely $f \in C_r(I)$ esetén (lásd Walsh [14, 351. oldal]) létezik egy $R_{m,n} \in \pi_{m,n}^c$, amelyre $\|f - R_{m,n}\|_{L_\infty(I)} = E_{m,n}^c(f)$. Nevezetesen, létezik egy $R_{1,1} \in \pi_{1,1}^c$, amelyre $\|x^2 - R_{1,1}\|_{L_\infty(I)} = E_{1,1}^c(x^2)$. Viszont (1.9)-ből világos, hogy $R_{1,1}(x)$ *nem lehet* valós minden valós x -re. Következésképpen, mivelhogy $x^2 - R_{1,1}(x)$ és komplex konjugáltja ugyanazon normával bírnak az I -n,

$$(1.10) \quad E_{1,1}^c(x^2) = \|x^2 - R_{1,1}(x)\|_{L_\infty(I)} = \|x^2 - \overline{R_{1,1}(x)}\|_{L_\infty(I)}.$$

Így $R_{1,1}$ és $\overline{R_{1,1}}$, amelyek mindketten *különböző* elemei a $\pi_{1,1}^c$ halmaznak, *mindketten* legjobb approximációi az x^2 függvénynek a $\pi_{1,1}^c$ halmazból I -n, és az x^2 függvény az (1.2) alatti mindkét tulajdonsággal rendelkezik az $m = n = 1$ esetben. Megjegyezzük, hogy ez az okoskodás általában is mutatja, hogy ha $E_{m,n}^c(f) < E_{m,n}^r(f)$, mint (1.2i)-ben, akkor ebből az (1.2ii) alatti nem-egyértelműség szükségszerűen következik.

A.A. Gonchar először 1986-ban említette meg a nem-egyértelműség lehetőségét a [2] cikkének egy lábjegyzetében. E lehetőséget K.N. Lungu, Gonchar egy diákja realizálta, aki 1971-ben [4] dolgozatában elégséges feltételeket adott az (1.2) tulajdonságok teljesülésére. Ettől függetlenül Saff és Varga [9,10] ugyanerre a felfedezésre jutott 1977-ben, általánosabb elégséges feltételeket kapva arra, hogy az $E_{m,n}^c(f) < E_{m,n}^r(f)$, ill. $E_{m,n}^c(f) = E_{m,n}^r(f)$ feltétel teljesüljön egy $f \in C_r(I)$ esetén. Az első esetben Saff és Varga elégséges feltételeit később Ruttan [6] élesítette, aki megmutatta, hogy $E_{m,n}^c(f) < E_{m,n}^r(f)$ teljesül, ha az f legjobb valós egyenletes approximációja a $\pi_{m,n}^r$ halmazból az I -n a maximális hibáját *semmilyen* $m + n + 1$ -nél hosszabb alternációs halmazon (lásd (1.4)) nem éri el, és hogy ez az alsó korlát általában a *lehető legjobb*. Hasonló eredmények áttekintése megtalálható a [12, Chapter 5] kiadványban.

Itt most a következő, Saff és Varga [10] által felvetett problémára koncentrálnunk. A nemnegatív egészek valamely (m, n) párjára definiáljuk a $\gamma_{m,n}$ nemnegatív valós számot a következőképpen:

$$(1.11) \quad \gamma_{m,n} := \inf\{E_{m,n}^c(f)/E_{m,n}^r(f) : f \in C_r(I) \setminus \pi_{m,n}^r\}.$$

A $\gamma_{m,n}$ szám meghatározása lényegében azt jelenti, hogy megvizsgáljuk, a *speciális* $C_r(I) \setminus \pi_{m,n}^r$ halmazbeli függvényeknek a $\pi_{m,n}^c$ halmazból vett legjobb egyenletes approximációja az I -n mennyivel *jobb*, mint a $\pi_{m,n}^r$ halmazból vett. Például az $x^2 \in C_r(I)$ függvényről ismeretes (lásd Bennett, Rudnick és Vaaler [1]), hogy $E_{1,1}^c(x^2) = (4/27)^{1/2} = 0,38490 \dots$, s így (1.6) miatt

$$\frac{E_{1,1}^c(x^2)}{E_{1,1}^r(x^2)} = 0,76980 \dots$$

Következésképpen ez az alábbi felső korlátot adja a $\gamma_{1,1}$ számra:

$$(1.12) \quad \gamma_{1,1} \leq 0,76980 \dots$$

Abból a célból, hogy a $\gamma_{m,0}$ számot bármely nemnegatív egész m esetén pontosan meghatározzuk, először is bebizonyítjuk az alábbi állítást.

1. ÁLLÍTÁS ([10]). Adott $f \in C_r(I)$ és a nemnegatív egészekből álló adott (m, n) számpár esetén

$$(1.13) \quad E_{m+n,2n}^r(f) \leq \inf_{g \in \pi_{m,n}^c} \|f - \operatorname{Re} g\|_{L_\infty(I)} \leq E_{m,n}^c(f) \leq E_{m,n}^r(f).$$

Bizonyítás. Mivel az (1.13) alatti utolsó egyenlőtlenség világos, tekintsünk egy tetszőleges $p_m/q_n \in \pi_{m,n}^c$ függvényt. Valós részeket véve nyilvánvaló, hogy

$$\left| f(x) - \frac{p_m(x)}{q_n(x)} \right| \geq \left| f(x) - \operatorname{Re} \left(\frac{p_m(x)}{q_n(x)} \right) \right| \quad (x \in I),$$

ami az (1.13) alatti második egyenlőtlenséget bizonyítja. Mivel $\operatorname{Re}(p_m/q_n) \in \pi_{m+n,2n}^r$, azért az (1.13) alatti első egyenlőtlenség is teljesül. \square

Az (1.13)-ban az $n = 0$ értéket véve láthatjuk, hogy

$$(1.14) \quad E_{m,0}^c(f) = E_{m,0}^r(f) \quad (f \in C_r(I); m = 0, 1, \dots),$$

s így (lásd (1.11))

$$(1.15) \quad \gamma_{m,0} = 1 \quad (m = 0, 1, \dots).$$

Kiderült, hogy az (1.11) alatti $\gamma_{m,n}$ konstans pontos meghatározása $m \geq 0$ és $n \geq 1$ esetén nehezebb, mint az (1.15) alatti $\gamma_{m,0}$ kiszámítása. A közelmúltban négy cikk foglalkozott a $\gamma_{m,n}$ viselkedésének leírásával $n \geq 1$ esetén. Először is Trefethen és Gutknecht [11] 1983-ban bebizonyította azt a meglehetősen figyelemreméltó eredményt, hogy

$$(1.16) \quad \gamma_{m,n} = 0 \quad (n \geq m + 3; m = 0, 1, \dots).$$

Ezután Levin [3] 1986-ban találta azt a kiegészítő eredményt, hogy

$$(1.17) \quad \gamma_{m,n} = 1/2 \quad (m + 1 \geq n \geq 1).$$

Levin bebizonyítása (1.17)-re abból áll, hogy egy direkt konstrukciót adott a $\gamma_{m,n} \leq 1/2$ bizonyítására, és egy algebrai módszert talált annak megmutatására, hogy $\gamma_{m,n} < 1/2$ lehetetlen $m + 1 \geq n \geq 1$ esetén. Az (1.16) és (1.17) alatti eredmények csak a $\gamma_{m,m+2}$ ($m \geq 0$) konstansok kérdését hagyták nyitva. Ezt az esetet legújabbban Ruttan és Varga intézte el 1989-ben, amikor is [7]-ben megmutatták, hogy $\gamma_{m,m+2} \leq 1/3$ és [8]-ban, hogy $\gamma_{m,m+2} < 1/3$ lehetetlen. Így

$$(1.18) \quad \gamma_{m,m+2} = 1/3 \quad (m = 0, 1, \dots).$$

m n	0	1	2	3	4	5	...
0	1	1	1	1
1	1/2	1/2	1/2	1/2
2	1/3	1/2	1/2	1/2
3	0	1/3	1/2	1/2
4	0	0	1/3	1/2
5	0	0	0	1/3	1/2
6	0	0	0	0
7
...

 5.1. táblázat: $\{\gamma_{m,n}\}_{m,n \geq 0}$.

Az (1.15) – (1.18) alatti eredményekből állítottuk össze az 5.1. Táblázat $\{\gamma_{m,n}\}_{m,n \geq 0}$ értékeit. Érdekes, hogy az (1.11) alatti $\{\gamma_{m,n}\}_{m,n \geq 0}$ konstansok csak a *négy* különböző 1, 1/2, 1/3 és 0 értéket veszik fel, és csak *három* különböző értéket, ha $n \geq 1$: 1/2, 1/3, és 0. Nevezetesen ebből következik, hogy bármely olyan ε esetén, amelyre $0 < \varepsilon < 1$, és bármely olyan (m, n) egész számokból álló számpárra, amelyre $m \geq 0$ és $n \geq 1$, létezik egy $f \in C_r(I) \setminus \pi_{m,n}^r$ úgy, hogy

$$(1.19) \quad \begin{cases} (i) & E_{m,n}^c(f) < (1 + \varepsilon)E_{m,n}^r(f)/2 \quad (m + 1 \geq n \geq 1); \\ (ii) & E_{m,m+2}^c(f) < (1 + \varepsilon)E_{m,m+2}^r(f)/3 \quad (m = 0, 1, \dots); \\ (iii) & E_{m,n}^c(f) < \varepsilon E_{m,n}^r(f) \quad (n \geq m + 3; m = 0, 1, \dots), \end{cases}$$

ami nem más, mint az (1.2i) egy élesebb formája.

A további paragrafusokban [8] alapján egy egységes eljárást ismertetünk a $\gamma_{m,n}$ konstansok meghatározására, midőn $m \geq 0$ and $n \geq 1$. Ugyancsak foglalkozunk az olyan $f \in C_r(I) \setminus \pi_{m,n}^r$ függvényekkel, amelyekre

$$(1.20) \quad \gamma_{m,n} < \frac{E_{m,n}^c(f)}{E_{m,n}^r(f)} < \gamma_{m,n} + \varepsilon \quad (m \geq 0; n \geq 1)$$

teljesül valamely adott $\varepsilon > 0$ mellett. Az ilyen, (1.20)-at kielégítő f -re, ha $g \in \pi_{m,n}^c$ és $h \in \pi_{m,n}^r$ olyanok, hogy

$$\|f - g\|_{L_\infty(I)} = E_{m,n}^c(f), \text{ és } \|f - h\|_{L_\infty(I)} = E_{m,n}^r(f),$$

úgyszintén érdekes kérdés a megfelelő

$$f(x) - g(x) \text{ és } f(x) - h(x) \quad (x \in I)$$

hibák viselkedése. E problémákból számos érdekes ábra adódik a fejezet hátralévő részében.

5.2. Felső korlátok a $\gamma_{m,n}$ konstansra

Az alábbi állítás segítségünkre lesz abban, hogy felső korlátokat kapjunk a $\gamma_{m,n}$ konstansokra $n \geq 1$ esetén.

2. ÁLLÍTÁS ([8]). *A nemnegatív egészek egy olyan adott (m, n) párja esetén, melyre $n \geq 1$, tegyük fel, hogy*

$$(2.1) \quad g \in \pi_{m,n}^c \setminus \pi_{m,n}^r, \quad \text{és} \quad \operatorname{Re} g \in C_r(I),$$

továbbá legyen $S \in C_r(I)$ olyan, hogy létezik $L \geq m + 2$ különböző $\{x_j\}_{j=1}^L$ pont úgy, hogy $-1 \leq x_1 < x_2 < \dots < x_L \leq 1$, és amelyre (rögzített $\lambda = 1$ vagy $\lambda = -1$ mellett)

$$(2.2) \quad \lambda(-1)^j [S(x_j) + \operatorname{Re} g(x_j)] > 0 \quad (j = 1, 2, \dots, L).$$

Ekkor

$$(2.3) \quad \gamma_{m,n} \leq \|S - i \operatorname{Im} g\|_{L_\infty(I)} / M,$$

ahol

$$(2.4) \quad M := \min_{1 \leq j \leq L} |S(x_j) + \operatorname{Re} g(x_j)|.$$

Bizonyítás. Legyen $f := S + \operatorname{Re} g$, úgyhogy $f \in C_r(I)$. Ha f -nek a $\pi_{m,n}^r$ halmazból vett legjobb approximációja azonosan zérus lenne, akkor az (1.5) alatti megállapodás szerint létezne egy I -beli, $\ell \geq m + 2$ hosszúságú alternációs halmaz. Mármost a (2.2) feltevés szerint f előjelben oszcillál az I intervallum $L \geq m + 2$ különböző pontjában, és ebből de la Vallée-Poussin tétele alapján (lásd [5, 162. oldal]) az alábbi alsó korlátot kapjuk az $E_{m,n}^r(f)$ mennyiségre:

$$E_{m,n}^r(f) \geq M,$$

ahol M a (2.4) alatt definiált mennyiség. Másrészt, mivel

$$E_{m,n}^c(f) \leq \|f - g\|_{L_\infty(I)} = \|S - i \operatorname{Im} g\|_{L_\infty(I)},$$

az (1.11) alatti $\gamma_{m,n}$ definíciójából adódik a (2.3) egyenlőtlenség. \square

Kiderült, hogy Trefethen és Gutknecht [11], Levin [3], valamint Ruttan és Varga [7] mindegyike légyegében a 2. Állítás egy variánsát használta, a $g(x)$ és $S(x)$ alkalmas választásával élve, a $\gamma_{m,n}$ korlátainak meghatározására. Azokat a speciális komplex racionális függvény konstrukciókat, amelyek a $\gamma_{m,n}$ éles felső korlátainak meghatározására vezettek, az alábbiakban ismertetjük.

Trefethen és Gutknecht [11] egy szellemes, komplex racionális függvény konstrukciójával kezdjük tárgyalásunkat, amiből következni fog, hogy $\gamma_{m,m+3} = 0$ ($m = 0, 1, \dots$). Tetszőleges, rögzített nemnegatív egész m és bármely olyan ε mellett, amelyre $0 < \varepsilon < 1$ (és amelyre $0 < \varepsilon < 1/(2m-1)$, ha $m \geq 1$), tekintsük az alábbi komplex racionális függvényt a $\pi_{m,m+3}^c \setminus \pi_{m,m+3}^r$ halmazból:

$$(2.5) \quad h_{m,\varepsilon}(x) := \frac{\varepsilon \prod_{j=1}^m [-1 + (2j-1)\varepsilon - x]}{[x+1+\varepsilon]^{m+1}(i\sqrt{\varepsilon}-x)(1+\varepsilon-x)},$$

(ahol, mint általában, $\prod_{j=1}^m := 1$, ha $m = 0$). A (2.5)-ből következik, hogy

$$(2.6) \quad \operatorname{Re} h_{m,\varepsilon}(x) = \frac{-\varepsilon x \prod_{j=1}^m [-1 + (2j-1)\varepsilon - x]}{[x+1+\varepsilon]^{m+1}(\varepsilon+x^2)(1+\varepsilon-x)} \in \pi_{m+1,m+4}^r,$$

és

$$(2.7) \quad \operatorname{Im} h_{m,\varepsilon}(x) = \frac{-\varepsilon\sqrt{\varepsilon} \prod_{j=1}^m [-1 + (2j-1)\varepsilon - x]}{[x+1+\varepsilon]^{m+1}(\varepsilon+x^2)(1+\varepsilon-x)} \in \pi_{m,m+4}^r.$$

A (2.6)-ból világos, hogy $\operatorname{Re} h_{m,\varepsilon}(x)$ -nak van $m+1$ különböző gyöke a $(-1, 0]$ intervallumban, m sűrűn elhelyezkedő gyöke a -1 jobb oldali környezetében, továbbá még egy gyöke az origóban. Ezek után definiáljunk $m+2$ különböző pontot, nevezetesen az $\{x_j(\varepsilon) := -1 + 2j\varepsilon\}_{j=0}^m$ és $x_{m+1}(\varepsilon) := 1$ pontokat, amelyek kielégítik a

$$-1 = x_0(\varepsilon) < x_1(\varepsilon) < \dots < x_{m+1}(\varepsilon) = 1$$

egyenlőtlenségeket. Ezen $m+2$ $\{x_j(\varepsilon)\}_{j=0}^{m+1}$ pont olyan, hogy bármely kettő között van pontosan egy gyöke a $\operatorname{Re} h_{m,\varepsilon}(x)$ függvénynek, és belátható, hogy $\operatorname{Re} h_{m,\varepsilon}(x)$ oszcillál előjel szerint e pontokban. Ezenkívül a $\operatorname{Re} h_{m,\varepsilon}(x) - 1 - \varepsilon$ pontbeli $m+1$ -edrendű, és az $1 + \varepsilon$ pontbeli elsőrendű pólusa azt eredményezi, hogy ezek az oszcillációk durván azonos abszolút értékűek, vagyis létezik egy az m -től függő, de az ε -tól független c konstans úgy, hogy (lásd [11])

$$(2.8) \quad (-1)^j \operatorname{Re} h_{m,\varepsilon}(x_j(\varepsilon)) \geq c \|\operatorname{Im} h_{m,\varepsilon}\|_{L_\infty(I)} / \sqrt{\varepsilon} \quad (j = 0, 1, \dots, m+1).$$

A (2.5)-beli $h_{m,\varepsilon}(x)$ racionális függvényre a 2. Állítás alapján kimondhatjuk az alábbi tételt:

3. TÉTEL (Trefethen és Gutknecht [11]). *Bármely nemnegatív egész m és bármely $n \geq m + 3$ egész esetén*

$$(2.9) \quad \gamma_{m,n} = 0 \quad (n \geq m + 3; m = 0, 1, \dots).$$

Bizonyítás. Először is megmutatjuk, hogy $\gamma_{m,m+3} = 0$ bármely nemnegatív egész m esetén. Az $n := m + 3$ jelöléssel, legyen $S := 0$, $L := m + 2$, és $g := h_{m,\varepsilon} \in \pi_{m,m+3}^c$, és alkalmazzuk a 2. Állítást. A fenti megfontolások mutatják, hogy a 2. Állításbeli (2.2) igaz, s (2.8)-ból látjuk (lásd (2.4)), hogy

$$M := \min_{0 \leq j \leq m+1} |\operatorname{Re} h_{m,\varepsilon}(x_j(\varepsilon))| \geq c \|\operatorname{Im} h_{m,\varepsilon}\|_{L_\infty(I)} / \sqrt{\varepsilon}.$$

A 2. Állítás (2.3) relációjából következik, hogy

$$(2.10) \quad \gamma_{m,m+3} \leq \sqrt{\varepsilon}/c.$$

Mivel (2.10)-ben c független ε -tól, és ε tetszőlegesen kicsi lehet, azért

$$(2.11) \quad \gamma_{m,m+3} = 0.$$

Továbbá, mivel $\pi_{m,n}^c \supset \pi_{m,m+3}^c$ bármely $n \geq m + 3$ esetén, ugyanaz a $h_{m,\varepsilon}$ függvény használható annak kikövetkeztetésére, hogy

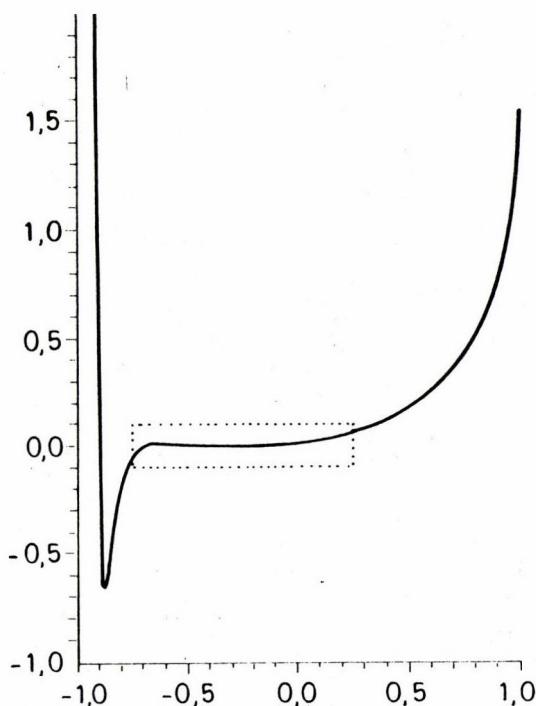
$$\gamma_{m,n} = 0 \quad (n \geq m + 3; m = 0, 1, \dots),$$

ami nem más, mint a kívánt (2.9) reláció. \square

A fenti konstrukcióban a (2.6) alatti $f(x) := \operatorname{Re} h_{m,\varepsilon}(x)$ egy olyan $C_r(I)$ halmazból függvény, amely szimultán approximálható a $\pi_{m,m+3}^r$ halmazhoz tartozó azonosan zérus függvénnyel, és a $\pi_{m,m+3}^c$ halmazbeli $h_{m,\varepsilon}(x)$ függvénnyel. Az f -nek az $m + 2$ számú különböző I -beli pontban bekövetkező, (2.8)-ban leírt oszcillációja miatt, a $\pi_{m,m+3}^r$ halmazbeli azonosan zérus függvény egy közel legjobb approximációja az f -nek a $\pi_{m,m+3}^r$ halmazból az I -n, és

$$E_{m,m+3}^r(f) \doteq \|f\|_{L_\infty(I)} = \|\operatorname{Re} h_{m,\varepsilon}\|_{L_\infty(I)}.$$

Az 5.1. ábrán az $f(x) := \operatorname{Re} h_{5,\varepsilon}(x)$ függvényt $x \in I$ esetén, $\varepsilon = 0,1$ mellett azért ábráztuk, hogy bemutassuk 7 oszcillációját az I -n. Mivel ezen oszcillációk némelyike nagyon kicsi, az 5.2 ábrán egy $(y\text{-ban})$ 15-szörös nagyításban adtuk meg az 5.1 ábra pontozott vonallal megjelölt részét, hogy e kicsiny oszcillációkat világosabban lássuk. Mármint ez esetben a komplex hiba definícióink szerint



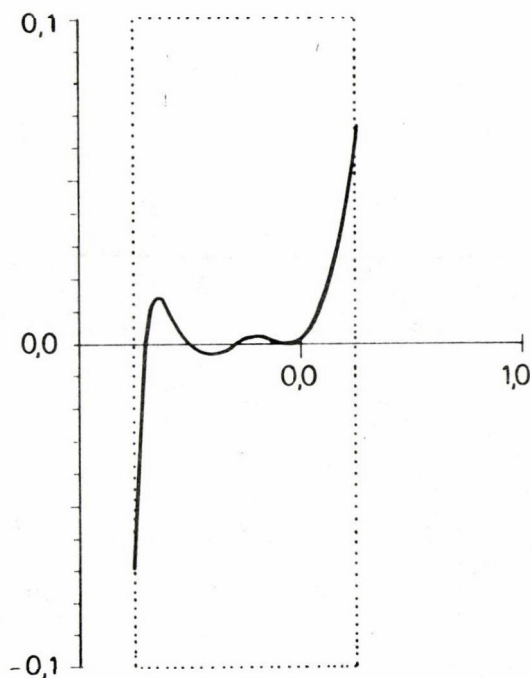
5.1. ábra: $\operatorname{Re} h_{5,\epsilon}(x)$.

$f(x) - g(x) = -i \operatorname{Im} h_{5,\epsilon}(x)$, és ennek nyomvonala a komplex síkon, midőn x növekszik -1 -től 1 -ig, érdektelen (s ezért elhagytuk), mivel e nyomvonal csak egy az imaginárius tengelyre szimmetrikus szakaszon történő mozgásra korlátozódik.

A 3. Tételbeli (2.9) eredménnyel kapcsolatban E. B. Saff rámutatott arra, hogy tetszőleges kis $\gamma_{m,n}$ számok létezése már Walsh [13, Theorem IV] egy 1934-es eredményéből is következik, bár Walsh ezen eredményének a $\gamma_{m,n}$ számokkal való kapcsolatát korábban senki sem vette észre. Nevezetesen, Walsh megmutatta, hogy bármely rögzített nemnegatív egész m esetén, az $\bigcup_{n=0}^{\infty} \pi_{m,n}^c$ függvények sűrűek a $C_c(I)$ halmazban, vagyis az összes folytonos, komplex értékű, I -n definiált függvények terén. Így, egy tetszőleges $f \in C_r(I) \subset C_c(I)$ függvényt választva, e sűrűség maga után vonja, hogy

$$(2.12) \quad \lim_{n \rightarrow \infty} E_{m,n}^c(f) = 0 \quad (m = 0, 1, \dots).$$

Másrészt, jelölje $T_j(x)$ ($j \geq 0$) a j -edik (elsőfajú) Csebisev-polinomot. Speciálisan, legyen $f(x) := T_{m+1}(x) \in C_r(I)$ és $r_{m,n} \equiv 0 \in \pi_{m,n}^c$, ekkor $f-0$ -nak létezik egy pontosan $m+2$ hosszúságú alternációs halmaza az I -ben, és az ismert $\|T_{m+1}\|_{L_\infty(I)} = 1$



5.2. ábra: Nagyítás.

relációból (lásd (1.4)) következik, hogy

$$(2.13) \quad E_{m,n}^r(f) = 1 \quad (n = 0, 1, \dots).$$

Mármost (2.12) és (2.13) kombinációja nyilvánvalóan azt adja, hogy

$$(2.14) \quad \lim_{n \rightarrow \infty} \gamma_{m,n} = 0 \quad (m = 0, 1, \dots).$$

Mindazonáltal láthatjuk, hogy a 3. Tételben szereplő (2.9) reláció a (2.14)-nek egy sokkal precízebb formája.

A következőkben Levin [3] konstrukcióját ismertetjük a $\gamma_{n+2k,n}$ felső korlátaira, ahol $n \geq 1$ és $k \geq 1$. (Később, az 5.3. pontban megmutatjuk, hogy ezek a felső korlátok élesek.)

Egy megfelelően kicsiny $\varepsilon > 0$ mellett tekintsük az alábbi komplex racionális függvényt:

$$(2.15) \quad h_{2k,n,\varepsilon}(x) := T_{2k}(x) \left(\frac{x - i\varepsilon}{x + i\varepsilon} \right)^n \in \pi_{2k+n,n}^c \setminus \pi_{2k+n,n}^r \quad (n \geq 1),$$

ahol $T_{2k}(x)$ ismét a $2k$ -adik (elsőfajú) Csebisev-polinomot jelöli, egy tetszőleges pozitív egész k esetén. Ekkor

$$(2.16) \quad \operatorname{Re} h_{2k,n,\varepsilon}(x) = T_{2k}(x) \operatorname{Re} \left\{ \left(\frac{x - i\varepsilon}{x + i\varepsilon} \right)^n \right\} \in \pi_{2n+2k,2n}^r.$$

Egy tetszőleges valós x esetén legyen $x + i\varepsilon = \rho e^{i\theta}$, úgyhogy $[(x - i\varepsilon)/(x + i\varepsilon)]^n = e^{-2ni\theta}$. Ebből adódik, hogy

$$(2.17) \quad \operatorname{Re} \{[(x - i\varepsilon)/(x + i\varepsilon)]^n\} = \cos(2n\theta),$$

ahonnan következik, hogy $\operatorname{Re} \{[(x - i\varepsilon)/(x + i\varepsilon)]^n\}$ rendelkezik $2n$ különböző gyökkel, nevezetesen $\left\{ x_j(\varepsilon) := -\varepsilon \cot \left[\frac{(2j+1)\pi}{4n} \right] \right\}_{j=0}^{2n-1}$, s ezek mind az I -beli origó egy ε -környezetében sűrűsödnek. A $T_{2k}(x)$ $2k$ különböző gyöke I -ben az $\{y_i := \cos \left[\frac{(2i+1)\pi}{4k} \right]\}_{i=1}^{2k-1}$ formulával adható meg. Mivel $T_{2k}(0) = (-1)^k \neq 0$, a $T_{2k}(x)$ ezen gyökei mind különböznek az $\{x_j(\varepsilon)\}_{j=0}^{2n-1}$ gyököktől elég kis $\varepsilon > 0$ esetén, s így $\operatorname{Re} h_{2k,n,\varepsilon}(x)$ összes, $2n + 2k$ számú (különböző) gyöke az I -ben megadható az

$$\{x_j(\varepsilon)\}_{j=0}^{2n-1} \cup \{y_i\}_{i=1}^{2k-1}$$

alakban, tetszőleges elég kis $\varepsilon > 0$ esetén. Ebből (2.16) és (2.17) segítségével nem nehéz kikövetkeztetni, hogy létezik $2n + 2k + 1$ különböző $\{w_j(\varepsilon)\}_{j=0}^{2n+2k}$ pont az I -ben, amelyek váltakozva közrefogják ezeket a gyököket, és amelyekben (lásd [3]) $\operatorname{Re} h_{2k,n,\varepsilon}(x)$ felveszi az $1 - o(1)$ értékeket váltakozó előjellel, vagyis

$$(2.18) \quad (-1)^j \operatorname{Re} h_{2k,n,\varepsilon}(w_j(\varepsilon)) \geq 1 - o(1) \quad (j = 0, 1, \dots, 2n + 2k),$$

midőn $\varepsilon \rightarrow 0$.

A (2.15) alatti $h_{2k,n,\varepsilon}$ racionális függvényről a 2. Állítás alapján az $m = 2k + n$ speciális esetben bebizonyítjuk a következő tételt.

4. TÉTEL (Levin [3]). A nemnegatív egészek egy tetszőleges olyan (m, n) párjára, amelyre $m + 1 \geq n \geq 1$, fennáll

$$(2.19) \quad \gamma_{m,n} \leq \frac{1}{2} \quad (m + 1 \geq n \geq 1).$$

Bizonyítás. Megmutatjuk, hogy $\gamma_{2k+n,n} \leq \frac{1}{2}$. Az $m := n + 2k$ mellett, ahol $k \geq 1$ és $n \geq 1$, legyen $S(x) := \operatorname{Re} h_{2k,n,\varepsilon}(x)$, $L := 2n + 2k + 1$, és $g(x) := h_{2k,n,\varepsilon}(x) \in \pi_{2k+n,n}^c$, és alkalmazzuk a 2. Állítást. Először is jegyezzük meg, hogy mivel a feltevés szerint $n \geq 1$, ezért $L \geq m + 2 = n + 2k + 2$. A fenti okoskodás szerint a 2. Állítás (2.2) relációja most is igaz, és (2.18)-ból láthatjuk, hogy (lásd (2.4))

$$M := \min_{0 \leq j \leq 2n+2k} |2 \operatorname{Re} h_{2k,n,\varepsilon}(w_j(\varepsilon))| \geq 2 - o(1), \text{ ha csak } \varepsilon \rightarrow 0.$$

Mármost számolással belátható, hogy

$$\begin{aligned} \|S(x) - i \operatorname{Im} g(x)\|_{L_\infty(I)} &= \|\operatorname{Re} h_{2k,n,\varepsilon}(x) - i \operatorname{Im} h_{2k,n,\varepsilon}(x)\|_{L_\infty(I)} \\ &= \|\overline{h_{2k,n,\varepsilon}(x)}\|_{L_\infty(I)} = 1, \end{aligned}$$

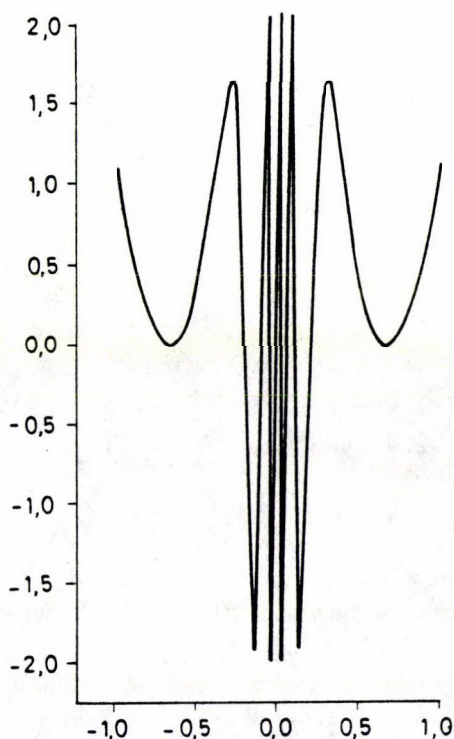
mivel (2.16) alapján $\left\| \left(\frac{x+i\varepsilon}{x-i\varepsilon} \right)^n \right\|_{L_\infty(I)} = 1 = \|T_{2k}\|_{L_\infty(I)}$ és $h_{2k,n,\varepsilon}(0) = 1$. Ezek után a 2. Állítás (2.3) relációja közvetlenül adja, hogy

$$(2.20) \quad \gamma_{2k+n,n} \leq \frac{1}{2 - o(1)} \quad (\varepsilon \rightarrow 0).$$

Itt az $\varepsilon \rightarrow 0$ határátmenettel élve,

$$\gamma_{2k+n,n} \leq \frac{1}{2} \quad (k = 1, 2, \dots; n \geq 1),$$

ami nem más, mint a (2.19) $m = 2k + n$ speciális esete. A fennmaradó esetekben a konstrukció (lásd [3]). \square

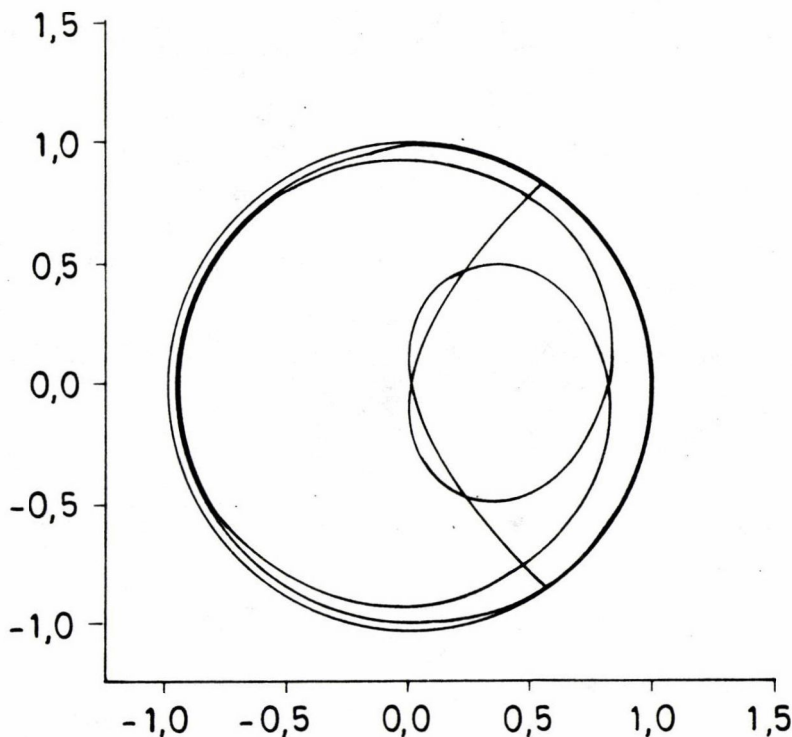


5.3. ábra: $2 \operatorname{Re} h_{2,5,\varepsilon}(x)$.

A fenti konstrukcióban az $f(x) = 2 \operatorname{Re} h_{2k,n,\varepsilon}(x) \in C_r(I)$ függvényt szimultán approximáltuk a $\pi_{2k+n,n}^r$ halmazbeli azonosan zérus és a $h_{2k,n,\varepsilon}(x) \in \pi_{2k+n,n}^c$ függvényekkel. Az 5.3. ábrán a $2 \operatorname{Re} h_{2,5,\varepsilon}(x)$ függvényt $x \in I$ esetén, az $\varepsilon = 0, 1$ értékkel ábrázoltuk annak érdekében, hogy megmutassuk 9 pontú közel ekvioszcillációját az I intervallumon. Ez esetben a komplex hiba pontosan

$$f(x) - g(x) = \overline{h_{2,5,\varepsilon}(x)} \quad (x \in I),$$

s ennek nyomvonalát a komplex síkon, midőn $x - 1$ -től 1-ig növekszik, az 5.4. ábrán szemléltettük. Érdekes megfigyelni, hogy e nyomvonal közel kör!



5.4. ábra: $\overline{h_{2,5,\varepsilon}(x)}$.

Végezetül ismertetjük Ruttan és Varga [7] konstrukcióját a $\gamma_{m,m+2} \leq 1/3$ felső korlátjának meghatározására bármely $m = 0, 1, \dots$ esetén. Egy tetszőleges $0 < \varepsilon < 1/(m+1)$ mellett legyen m valamely rögzített nemnegatív páros egész, és tekintsük a következő függvényeket:

$$(2.21) \quad \ell_j(z) := \ell_j(z; \varepsilon, m) := \frac{-\frac{2\varepsilon i}{3}(-1)^j}{z - 1 + \frac{2j}{m+1} - \varepsilon i} \quad (j = 0, 1, \dots, m+1).$$

A (2.21)-ből nyilvánvaló, hogy

$$(2.22) \quad \ell_j \left(1 - \frac{2j}{m+1} \right) = \frac{2}{3}(-1)^j, \text{ és } \ell_j \left(1 - \frac{2j}{m+1} \pm \varepsilon \right) = \frac{(1 \mp i)(-1)^j}{3},$$

bármely $j = 0, 1, \dots, m+1$ esetén. Mivel mindegyik $\ell_j(z)$ egy Möbius transzformáció, azért minden egyes ℓ_j leképezi az \mathbb{R} valós tengelyt a komplex sík valamely általánosított körére. Az $\ell_j(\infty) = 0$ miatt ez az általánosított kör szükségszerűen keresztülmegy az origón. Továbbá, mivel az $\ell_j(z)$ függvény $1 - (2j)/(m+1) + \varepsilon i$ pólusának az \mathbb{R} -re vonatkozó tükrösképe $w_j := 1 - (2j)/(m+1) - \varepsilon i$, a (2.21)-ből következik, hogy

$$(2.23) \quad \ell_j(w_j) = \frac{1}{3}(-1)^j \quad (j = 0, 1, \dots, m+1).$$

Így a Möbius transzformációkra vonatkozó szimmetria-elv alapján az \mathbb{R} halmaz a $w = \ell_j(z)$ leképezéssel átmegy a $\frac{1}{3}(-1)^j$ középpontú és $1/3$ sugarú körbe (mivel ez a kör átmegy az origón). Így geometriailag világos, hogy

$$(2.24) \quad \|\operatorname{Re} \ell_j\|_{L_\infty(\mathbb{R})} = \frac{2}{3}, \text{ és } \|\operatorname{Im} \ell_j\|_{L_\infty(\mathbb{R})} = \frac{1}{3},$$

hacsak $j = 0, 1, \dots, m+1$.

Ahhoz, hogy kiterjesszük a (2.24) alatti eredményt, definiáljuk az $I_k(m)$ valós intervallumokat az

$$(2.25) \quad I_k(m) := \left[1 - \frac{2k+1}{m+1}, 1 - \frac{2k-1}{m+1} \right] \cap I \quad (k = 0, 1, \dots, m+1)$$

formulákkal, úgyhogy ezen intervallumok lefedik az $I := [-1, +1]$ intervallumot, vagyis

$$\bigcup_{k=0}^{m+1} I_k(m) = I.$$

Az $\ell_j(x)$ és $I_k(m)$ definíciójából következik, hogy (mivel m rögzített)

$$(2.26) \quad \|\ell_j\|_{L_\infty(I_k(m))} = O(\varepsilon) \text{ bármely } k \neq j \text{ esetén (midőn } \varepsilon \rightarrow 0),$$

továbbá (2.22) alapján

$$(2.27) \quad \|\operatorname{Re} \ell_j\|_{L_\infty(I_j(m))} = \frac{2}{3}, \text{ és } \|\operatorname{Im} \ell_j\|_{L_\infty(I_j(m))} = \frac{1}{3} \quad (j = 0, 1, \dots, m+1).$$

Mármost tekintsük a

$$(2.28) \quad h(x) = h(x; \varepsilon, m) := \sum_{j=0}^{m+1} \ell_j(x)$$

komplex racionális függvényt. Ennek racionális törtfüggvény alakja

$$(2.29) \quad h(x) = \frac{\frac{-2\epsilon i}{3} \sum_{j=0}^{m+1} (-1)^j \prod_{\substack{k=0 \\ k \neq j}}^{m+1} \left\{ x - 1 + \frac{2k}{m+1} - \epsilon i \right\}}{\prod_{k=0}^{m+1} \left\{ x - 1 + \frac{2k}{m+1} - \epsilon i \right\}},$$

s ezért h a $\pi_{m+1, m+2}^c$ halmazhoz tartozik. Továbbá (2.29)-ben a $h(x)$ számlálója

$$\frac{-2\epsilon i}{3} \left\{ x^{m+1} \sum_{j=0}^{m+1} (-1)^j + \text{alacsonyabb fokú tagok } (0 \leq s \leq m) \right\}.$$

Mivel m a feltevés szerint páros, ezért $\sum_{j=0}^{m+1} (-1)^j = 0$, amiből következik, hogy $h \in \pi_{m, m+2}^c$. Pontosabban, bebizonyítható, hogy a (2.29)-ben a $h(x)$ számlálójában az x^m együtthatója

$$-\frac{2(m+2)\epsilon i}{3(m+1)} \neq 0$$

s így h eleme a $\pi_{m, m+2}^c$ halmaznak, de nem eleme a $\pi_{s, m+2}^c$ -nek, tetszőleges $s < m$ esetén. Itt érdemes megemlíteni, hogy (2.28) nem más, mint $h(x)$ *parciális törtekre bontása*.

A (2.29) alatti h racionális függvény és a 2. Állítás segítségével bebizonyítjuk az alábbi tételt.

5. TÉTEL ([7]). *Tetszőleges nemnegatív egész m esetén*

$$(2.30) \quad \gamma_{m, m+2} \leq \frac{1}{3}.$$

Bizonyítás. Tetszőleges rögzített nemnegatív páros m egész esetén tekintsük az \mathbb{R} halmazon definiált

$$(2.31) \quad s(u) := \begin{cases} \frac{1-u^2}{1+u^2}, & -1 \leq u \leq +1, \\ 0, & \text{egyébként} \end{cases}$$

függvényt. Ekkor tehát $s(0) = 1$, $s(\pm 1) = 0$, és $0 < s(u) < 1$ bármely $0 < |u| < 1$ esetén. Figyelembevéve, hogy $0 < \epsilon < 1/(m+1)$, legyen

$$(2.32) \quad S_{\epsilon, m}(x) := \frac{1}{3} \sum_{j=0}^{m+1} (-1)^j s\left(\frac{x-1+\frac{2j}{m+1}}{\epsilon^2}\right) \quad (x \in \mathbb{R}).$$

Ebből következik, hogy $S_{\varepsilon,m}(x)$ egy valós, folytonos függvény az \mathbb{R} halmazon, továbbá

$$(2.33) \quad S_{\varepsilon,m} \left(1 - \frac{2j}{m+1} \right) = \frac{(-1)^j}{3}, \text{ és } S_{\varepsilon,m} \left(1 - \frac{2j}{m+1} \pm \varepsilon^2 \right) = 0 \\ (j = 0, 1, \dots, m+1).$$

Megjegyezzük, hogy geometriailag $S_{\varepsilon,m}(x)$ $m+2$ alternáló csúccsal rendelkezik az I intervallumon, mégpedig minden egyes $\left[1 - \frac{2j}{m+1} - \varepsilon^2, 1 - \frac{2j}{m+1} + \varepsilon^2 \right]$ ($j = 0, 1, \dots, m+1$) intervallumon eggyel (ezen intervallumok diszjunktak). A fenti definíciókkal legyen $n := m+2$, $L := m+2$, $S(x) := S_{\varepsilon,m}(x)$ (lásd (2.32)), $g(x) := h(x; \varepsilon, m)$ (lásd (2.28)), és alkalmazzuk a 2. Állítást. Ilymódon (2.22), (2.26) – (2.28) és (2.33) alapján azt kapjuk, hogy

$$(2.34) \quad (-1)^j \left[S_{\varepsilon,m} \left(1 - \frac{2j}{m+1} \right) + \operatorname{Re} g \left(1 - \frac{2j}{m+1} \right) \right] = 1 + O(\varepsilon) \\ (j = 0, 1, \dots, m+1),$$

midőn $\varepsilon \rightarrow 0$, úgyhogy (lásd (2.4))

$$(2.35) \quad M := \min_{0 \leq j \leq m+1} \left| S_{\varepsilon,m} \left(1 - \frac{2j}{m+1} \right) + \operatorname{Re} g \left(1 - \frac{2j}{m+1} \right) \right| = 1 + O(\varepsilon).$$

Másrészt tekintsük az $S_{\varepsilon,m}(x) - i \operatorname{Im} g(x)$ függvényt az I intervallumon. A (2.25)-beli $I_k(m)$ intervallumra (2.26) alapján következik, hogy

$$S_{\varepsilon,m}(x) - i \operatorname{Im} g(x) = S_{\varepsilon,m}(x) - i \operatorname{Im} \ell_k(x) + O(\varepsilon) \quad (x \in I_k(m)).$$

Továbbá egy rövid számolással (2.27) és (2.32) azt mutatja, hogy

$$\|S_{\varepsilon,m}(x) - i \operatorname{Im} \ell_k(x)\|_{L_\infty(I_k(m))} = \frac{1}{3} + O(\varepsilon) \quad (k = 0, 1, \dots, m+1),$$

s így (2.26) miatt

$$(2.36) \quad \|S_{\varepsilon,m}(x) - i \operatorname{Im} g(x)\|_{L_\infty(I)} = \frac{1}{3} + O(\varepsilon) \quad (\text{midőn } \varepsilon \rightarrow 0).$$

Ezért a 2. Állítás (2.3) relációjából következik, hogy

$$\gamma_{m,m+2} \leq \frac{1}{3} + O(\varepsilon) \quad (\text{midőn } \varepsilon \rightarrow 0),$$

és innen az $\varepsilon \rightarrow 0$ határátmenettel kapjuk, hogy

$$\gamma_{m,m+2} \leq \frac{1}{3},$$

ami a kívánt (2.30) alatti eredmény, amikor m páros.

Abban az esetben, ha m páratlan, egyszerűen módosítjuk a (2.21) definíciót az

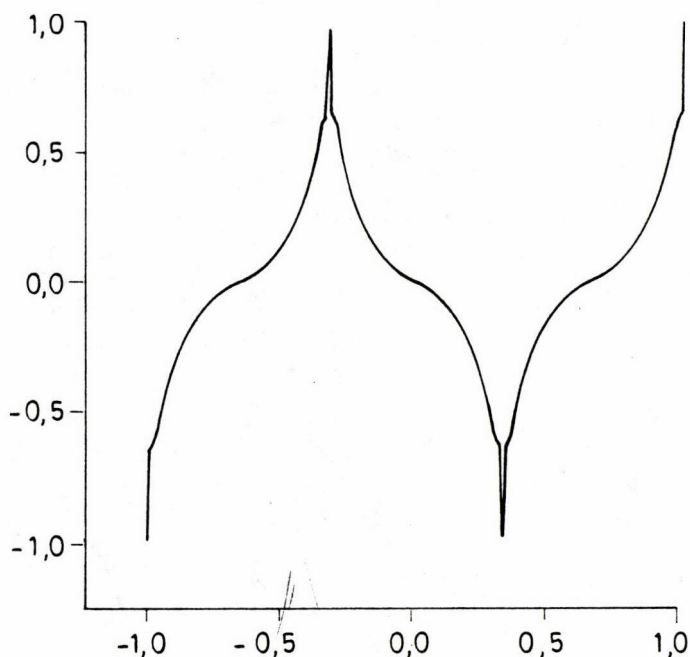
$$(2.37) \quad \ell_j(z) = \ell_j(z; \varepsilon, m) := \frac{\frac{-2\varepsilon i}{3} \mu_j (-1)^j}{z - 1 + \frac{2j}{m+1} - \varepsilon \mu_j i} \quad (j = 0, 1, \dots, m+1)$$

formulára, ahol a $\{\mu_j\}_{j=0}^{m+1}$ sorozat $m+2$ tetszőleges rögzített pozitív számból áll, amelyek kielégítik a $0 < \mu_j < 1$ egyenlőtlenséget, továbbá

$$(2.38) \quad \sum_{j=0}^{m+1} (-1)^j \mu_j = 0 \text{ és } \sum_{j=0}^{m+1} j(-1)^j \mu_j \neq 0.$$

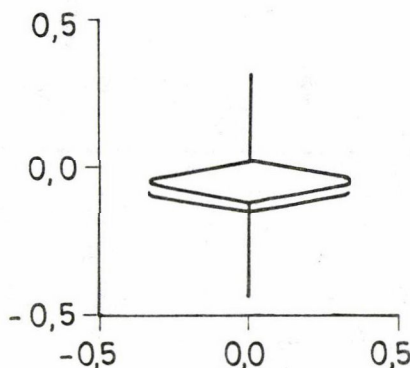
Ismét csak a 2. Állítást alkalmazva adódik a kívánt eredmény. \square

A fenti konstrukcióban $f_{\varepsilon, 2m}(x) := S_{\varepsilon, 2m}(x) + \operatorname{Re} h(x; \varepsilon, 2m) \in C_r(I)$ egy olyan függvény, amelyet szimultán approximáltunk a $\pi_{2m, 2m+2}^r$ halmazhoz tartozó azonosan zérus függvénnyel, valamint a $h(x; \varepsilon, 2m) \in \pi_{2m, 2m+2}^c$ függvénnyel. Az 5.5. ábrán az $S_{\varepsilon, 2}(x) + \operatorname{Re} h(x; \varepsilon, 2)$ függvényt $x \in I$ esetén az $\varepsilon = 0,1$ érték mellett ábrázoltuk, hogy megmutassuk négy csúcsát és négy közel ekvioszcillációs pontját az I intervallumon.



5.5. ábra: $S_{\varepsilon, 2}(x) + \operatorname{Re} h_{x; \varepsilon, 2}$.

Ez esetben a komplex hiba $S_{\varepsilon,2}(x) - i \operatorname{Im} h(x; \varepsilon, 2)$, amelynek nyomvonalát a komplex síkon, midőn $x - 1$ -től 1 -ig nő, az 5.6 ábrán szemléltettük. A (2.36) alapján ez a nyomvonal a $\{z \in \mathbb{C} : |z| \leq \frac{1}{3} + O(\varepsilon)\}$ körlemezre korlátozódik, de nem hasonlít egy körre.



5.6. ábra: $S_{\varepsilon,2} - i \operatorname{Im} h(x; \varepsilon, 2)$.

A jelen részbeli konstrukciók és ábrák együttesen azt mutatják, hogy a $C_r(I)$ halmazbeli függvényeket mennyire *különbözőképpen* kell választani ahhoz, hogy a $\gamma_{m,n}$ számra éles felső korlátokat kapjunk a (2.9), (2.19) és (2.30) esetekben.

5.3. A $\gamma_{m,n}$ szám alsó korlátai

A $\gamma_{m,n}$ szám alsó korlátainak meghatározásához először is felidézzük Ruttan és Varga [8, 4. és 5. Tétel] alábbi két eredményét.

Egy adott valós vagy komplex p polinom esetén jelölje ∂p a p pontos fokszámát. Ha $R = p/q$ folytonos az I -n, ahol p és q valós polinomok, akkor világos, hogy $\operatorname{Re} R = R$ legfeljebb ∂p alkalommal válthat előjelet az I -n, mivel az R minden egyes jelváltása megfelel a p egy gyökének. Viszont, ha $R = p/q$ egy folytonos, komplex értékű függvény az I -n, akkor a $\operatorname{Re} R$ előjelváltásainak lehetséges száma az I -n nemcsak ∂p -tól és ∂q -tól függ, hanem a $\operatorname{Re} R$ oszcillációinak nagyságrendjétől is az I -n. Ez a témája a [8] dolgozat alábbi eredményének. Vezessük még be a $[x]$ jelölést azon legnagyobb N egész számra, amelyre $N \leq x$.

6. TÉTEL([8]). Legyen $\phi = p/q$ egy olyan komplex racionális függvény, amelynek nincsenek pólusai az I -ben, és amely kielégíti az $\|\operatorname{Im} \phi\|_{L_\infty(I)} \leq 1$ egyenlőtlenséget. Tegyük fel, hogy léteznek $d > 0$ és $\{x_j\}_{j=1}^L$ valós számok úgy, hogy $-1 \leq x_1 < x_2 < \dots < x_L \leq 1$, és amelyekre (egy rögzített $\lambda = 1$ vagy $\lambda = -1$ értékkel)

$$(3.1) \quad \lambda(-1)^j \operatorname{Re} \phi(x_j) \geq d \quad (j = 1, 2, \dots, L).$$

Ha $\partial q \leq \partial p$ és $d \geq 1$, akkor

$$(3.2) \quad L \leq \partial p + 1.$$

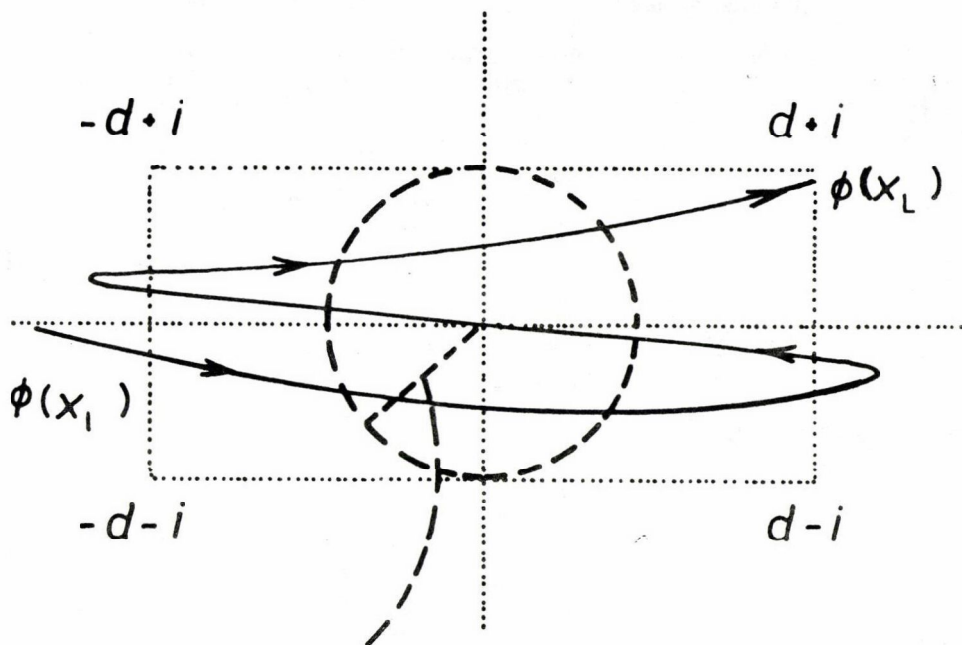
Hasonlóképpen, ha $\partial q > \partial p$, akkor

$$(3.3) \quad L \leq \partial q, \text{ valahányszor } d \geq 1,$$

és

$$(3.4) \quad L \leq \left\lfloor \frac{\partial p + \partial q + 1}{2} \right\rfloor \text{ valahányszor } d \geq 2.$$

Bizonyítás. A (3.2)-t egy Levin [3] dolgozatában kidolgozott geometriai módszer segítségével bizonyítjuk. Feltéve, hogy $\partial q \leq \partial p$ és $d \geq 1$, legyen B az a zárt téglalap a komplex síkon, amelynek csúcsai $\pm d \pm i$, úgyhogy a $C := \{z : |z| = 1\}$ kör részhalmaza a B -nek, amint az az 5.7. ábrán látható. A (3.1) feltétel és $\|\operatorname{Im} \phi\|_L \leq 1$ maga után vonja, hogy a $\Gamma_1 := \{z = \phi(x) : x \in \mathbb{R}\}$ görbe (a kiterjesztett komplex számsíkon) metszi a B függőleges oldalait, és így a C -t is, mégpedig $2(L-1)$ pontban, míg x növekszik x_1 -től x_L -ig. (itt azon pontokat, ahol Γ_1 érintőlegesen C -vel, kétszeresen számoltuk.) Ha $\phi(t), t \in I$ esetén egy ilyen metszéspontja Γ_1 -nek és C -nek, akkor



5.7. ábra: A B téglalap.

$$|\phi(t)|^2 = \left| \frac{p(t)}{q(t)} \right|^2 = 1$$

s így t gyöke a valós

$$P(x) := |p(x)|^2 - |q(x)|^2$$

polinomnak. E megfontolások mutatják, hogy a $P(x)$ -nek legalább $2(L-1)$ gyöke van az I -ben. Mivel $\partial p \geq \partial q$, ezért $\partial P \leq 2\partial p$. Így $2(L-1) \leq \partial P \leq 2\partial p$, amiből $L \leq \partial p + 1$, vagyis a kívánt (3.2) egyenlőtlenség következik. Az 5. Tétel további állításainak bizonyítása hasonlóképpen geometriai természetű, de sok eset megkülönböztetését igényli (lásd [8]). \square

Megjegyezzük, hogy mint azt [8]-ban megmutattuk, a 6. Tétel (3.2) – (3.4) eredményei *élesek*, az alábbi értelemben: i) léteznek a megfelelő feltételeket kielégítő komplex racionális függvények úgy, hogy az L -re vonatkozó, (3.2) – (3.4) alatti felső korlátok eléretnek, vagyis (3.2) – (3.4)-ben lehetséges az egyenlőség; és ii) tetszőleges pozitív $d < 1$ (< 2) esetén léteznek olyan racionális függvények, amelyek a 6. Tétel összes feltételeit kielégítik, kivéve a (3.3) (ill. (3.4)) alatti, d -re vonatkozó feltételeket, és amelyek a (3.3) (ill. (3.4)) alatti, L -re vonatkozó korlátokat meghaladják.

A 6. Tétel segítségével az alábbi, $\gamma_{m,n}$ -ra vonatkozó alsó korlátok kaphatók.

7. TÉTEL ([8]). Legyen (m, n) a nemnegatív egészek egy olyan párja, amelyre $n \geq 1$, legyen $f \in C^r(I) \setminus \pi_{m,n}^r$, és legyenek $r_{m,n}$ ill. $R_{m,n}$ az f legjobb egyenletes racionális approximációi a $\pi_{m,n}^r$ ill. $\pi_{m,n}^c$ halmazból az I -n. Ekkor

$$(3.5) \quad \frac{\|f - R_{m,n}\|_{L_\infty(I)}}{\|f - r_{m,n}\|_{L_\infty(I)}} > \frac{1}{2} \quad \text{ha } m+1 \geq n \geq 1,$$

és

$$(3.6) \quad \frac{\|f - R_{m,n}\|_{L_\infty(I)}}{\|f - r_{m,n}\|_{L_\infty(I)}} > \frac{1}{3} \quad \text{ha } m+2 \geq n \geq 1.$$

Következésképpen

$$(3.7) \quad \gamma_{m,n} = \frac{1}{2}, \quad \text{ha } m+1 \geq n \geq 1,$$

és

$$(3.8) \quad \gamma_{m,m+2} = 1/3.$$

Bizonyítás. Legyen $s := \|f - R_{m,n}\|_{L_\infty(I)} / \|f - r_{m,n}\|_{L_\infty(I)}$, úgyhogy $0 \leq s \leq 1$, továbbá legyen $e := f - r_{m,n}$, $R_{m,n} := p_1/q_1$ és $r_{m,n} := p_2/q_2$, ahol a (p_1, q_1) és

(p_2, q_2) párokról feltesszük, hogy nincs közös faktoruk. Mivel $f \notin \pi_{m,n}^r$, feltételezhetjük, hogy alkalmas konstanssal való szorzás után $\|e\|_{L_\infty(I)} = 1$, vagyis

$$(3.9) \quad s = \|f - R_{m,n}\|_{L_\infty(I)}.$$

Mivel $r_{m,n}$ az f legjobb egyenletes approximációja a $\pi_{m,n}^r$ halmazból az I -n, létezik (lásd (1.5)) legalább

$$(3.10) \quad L \geq 2 + \max[m + \partial q_2; n + \partial p_2] = m + m + 2 - \min[m - \partial p_2; n - \partial q_2]$$

különböző $\{x_j\}_{j=1}^L$ pont úgy, hogy $-1 \leq x_1 < x_2 < \dots < x_L \leq 1$, és $e(x_j) = (-1)^j \lambda$ minden $1 \leq j \leq L$ esetén (rögzített $\lambda = 1$ vagy $\lambda = -1$ mellett). Szükség esetén ismét -1 -gyel szorozva feltehetjük, hogy $\lambda = 1$, vagyis $e(x_1) = -1$.

E normalizációkkal élve kapjuk, hogy

$$s = \|f - R_{m,n}\|_{L_\infty(I)} \geq |f(x_j) - R_{m,n}(x_j)| = |(-1)^j + r_{m,n}(x_j) - R_{m,n}(x_j)|$$

bármely $1 \leq j \leq L$ esetén, ahonnan a valós részekre következik, hogy

$$(3.11) \quad (-1)^j \{\operatorname{Re} R_{m,n}(x_j) - r_{m,n}(x_j)\} \geq 1 - s \quad (j = 1, 2, \dots, L).$$

Mármost legyen $\phi(x) := (R_{m,n}(x) - r_{m,n}(x))/s = p(x)/q(x)$, ahol p és q komplex, közös faktor nélküli polinomok. A (3.11)-ből következik, hogy

$$(3.12) \quad (-1)^j \operatorname{Re} \phi(x_j) \geq (1 - s)/s =: d \quad (j = 1, 2, \dots, L),$$

és mivel (lásd (3.9)) $s = \|f - R_{m,n}\|_{L_\infty(I)} \leq \|\operatorname{Im} R_{m,n}\|_{L_\infty(I)}$, hasonlóképpen adódik, hogy

$$(3.13) \quad \|\operatorname{Im} \phi\|_{L_\infty(I)} \leq 1.$$

Most tekintsük a 7. Tétel (3.5) állítását, ami a fenti redukciókkal és definíciókkal élve nem más, mint hogy

$$m + 1 \geq n \geq 1 \text{ maga után, hogy } s > 1/2,$$

vagy másképp kifejezve $s \leq 1/2$ maga után vonja, hogy $m + 1 < n$, azaz

$$(3.14) \quad s \leq 1/2 \text{ maga után vonja, hogy } m + 2 \leq n.$$

Ez utóbbi bebizonyítására tegyük fel, hogy $0 < s \leq 1/2$, amiből (3.12) alapján következik, hogy $d \geq 1$. Mivel a 6. Tétel összes feltétele teljesül, a (3.2) és (3.3)-ból következik, hogy

$$(3.15) \quad L \leq \partial p + 1, \text{ ha } \partial q \leq \partial p, \text{ és } L \leq \partial q, \text{ ha } \partial q > \partial p.$$

Feltéve, hogy $\partial q > \partial p$, a (3.15) második része azt adja, hogy $L \leq \partial q$, ami a (3.10) és a $\phi = p/q = (R_{m,n} - r_{m,n})/s = (p_1/q_1 - p_2/q_2)/s$ definíció segítségével úgy fejezhető ki, hogy

$$m + m + 2 - \min[m - \partial p_2; n - \partial q_2] \leq \partial q = \partial q_1 + \partial q_2,$$

vagy ezzel ekvivalens módon

$$(3.16) \quad \{n - \partial q_1\} + \{(n - \partial q_2) - \min[m - \partial p_2; n - \partial q_2]\} \leq n - (m + 2).$$

Mivel (3.16)-ban a két kapcsos zárójeles tag nyilvánvalóan nemnegatív, ezért $n - (m + 2) \geq 0$, vagyis $m + 2 \leq n$, ami a (3.14) alatti állítás azon feltétellel, hogy $\partial q > \partial p$. Hasonló okoskodás mutatja, hogy a $\partial q \leq \partial p$ feltétel ellentmondásra vezet, s ezért a 7. Tétel (3.5) állítása igaz. Viszont ekkor (3.5) alapján

$$\gamma_{m,n} \geq \frac{1}{2}, \quad \text{ha } m + 1 \geq n \geq 1,$$

s mivel (2.19) alapján a fordított egyenlőtlenség ugyancsak teljesül, azért (Levin [3] okoskodását követve)

$$\gamma_{m,n} = \frac{1}{2}, \quad \text{ha } m + 1 \geq n \geq 1,$$

ami nem más, mint (3.7). A (3.6) hasonlóképpen látható be (lásd [8]), amiből (3.20) segítségével következik (3.8). \square

Megjegyezzük, hogy (3.5)-ben a szigorú egyenlőtlenség azt az erősebb eredményt adja, hogy *nincs* olyan $f \in C_r(I) \setminus \pi_{m,n}^r$, amelyre

$$(3.17) \quad \frac{\|f - R_{m,n}\|_{L_\infty(I)}}{\|f - r_{m,n}\|_{L_\infty(I)}} = \gamma_{m,n} = \frac{1}{2}, \quad \text{ha } m + 1 \geq n \geq 1,$$

ahol $r_{m,n} \in \pi_{m,n}^r$ ill. $R_{m,n} \in \pi_{m,n}^c$ az f legjobb egyenletes approximációi a $\pi_{m,n}^r$, ill. $\pi_{m,n}^c$ halmazból az I -n. Így $\gamma_{m,n}$ egy valódi infimum és nem minimum), mint azt (1.11)-ben definiáltuk, hacsak $m + 1 \geq n \geq 1$. Ugyanez hasonló okokból igaz $\gamma_{m,m+2} = 1/3$ -ra is, hacsak $m \geq 0$, és $\gamma_{m,n} = 0$ -ra, ha $n \geq m + 3$. Ezért függött az 5.2. pont valamennyi konstrukció egy kis $\varepsilon > 0$ paramétertől.

IRODALOM

- [1] C. BENNETT, K. RUDNICK és J.D. VAALER, „Best uniform approximation by linear fractional transformations”, *J. Approx. Theory* **25** (1979), 204–224.
- [2] A.A. GONCHAR, „The rate of approximation by rational fractions and the properties of functions, (oroszul)”, *Proceedings of the International Congress of Mathematicians* (Moscow), 1966 (Izdat. „Mir”, Moscow, 1968), 329–356.

- [3] A. LEVIN, „On the degree of complex rational approximation to real functions”, *Constr. Approx.* **2** (1986), 213–219.
- [4] K.N. LUNGU, „Best approximation by rational functions”, *Mat. Zametki* **10** (1971), 11–15, (oroszul).
- [5] D. MEINARDUS, *Approximation of Functions: Theory and Numerical Methods* (Springer-Verlag Inc., New York, 1967).
- [6] A. RUTTAN, „The length of the alternation set as a factor in determining when a best real rational approximation is also a best complex rational approximation”, *J. Approx. Theory* **31** (1981), 230–243.
- [7] A. RUTTAN és R.S. VARGA, „Real vs. complex rational Chebyshev approximation on an interval: $\gamma_{m,m+2} \leq 1/3$ ”, *Rocky Mountain J. of Math.* **19** (1989), 375–381.
- [8] A. RUTTAN és R.S. VARGA, „A unified theory for real vs. complex rational Chebyshev approximation on an interval”, *Trans. Amer. Math. Soc.* **312** (1989), 681–697.
- [9] E.B. SAFF és R.S. VARGA, „Nonuniqueness of best approximating complex rational functions”, *Bull. Amer. Math. Soc.* **83** (1977), 375–377.
- [10] E.B. SAFF és R.S. VARGA, „Nonuniqueness of best complex rational approximants to real functions on real intervals”, *J. Approx. Theory* **23** (1978), 78–85.
- [11] L.N. TREFETHEN és M.H. GUTKNECHT, „Real vs. complex rational Chebyshev approximation on an interval”, *Trans. Amer. Math. Soc.* **280** (1983), 555–561.
- [12] R.S. VARGA, *Topics in Polynomial and Rational Interpolation and Approximation* (University of Montreal Press, Montreal, 1982).
- [13] J.L. WALSH, „On approximation to an analytic function by rational functions of best approximations”, *Math. Zeit.* **38** (1934), 163–176.
- [14] J.L. WALSH, *Interpolation and Approximation by Rational Functions in the Complex Domain*, 5-th ed, Colloq. Publ. Vol. 20 (Amer. Math. Soc., Providence, R.I., 1969).

6. FEJEZET

A Jensen egyenlőtlenség általánosításai
alacsonyfokú koncentrációjú polinomokra

6.1. Alacsonyfokú koncentrációjú polinomok

A közelmúltban B. Beauzamy és P. Enflo [2] egy új, polinomokkal kapcsolatos elnevezést, nevezetesen az *alacsonyfokú koncentrációjú* polinomok fogalmát vezette be. A jelen fejezetben újabb eredményeket mutatunk be e fogalom alkalmazásaképpen a klasszikus analízisbeli Jensen egyenlőtlenség általánosításaira. Érdemes megemlíteni, hogy több, itt tárgyalandó eredményt olyan szimbolikus számítástechnikai eljárásokat használó numerikus megfontolások motiváltak, mint pl. a MACSYMA (lásd Beauzamy [3]).

Először is legyen $p(z) = \sum_{j=0}^m a_j z^j$ egy tetszőleges komplex polinom ($\neq 0$). Megadva egy tetszőleges valós d számot a $(0,1)$ intervallumban, valamint egy k nemnegatív egészet, azt fogjuk mondani, hogy $p(z)$ *k -adfokú koncentrációja* d , ha

$$(1.1) \quad \sum_{j=0}^k |a_j| \geq d \sum_{j=0}^m |a_j|.$$

(Mint azt később a 6.3. pontban látni fogjuk, ez a fogalom kiterjeszthető más függvényekre is.)

A témában elért első eredmény a következő.

1. TÉTEL (Beauzamy és Enflo [2], [3]). Adott egy tetszőleges valós d szám a $(0,1)$ intervallumban, valamint egy nemnegatív egész k szám, definiáljuk a csak d -től és k -tól függő $\tilde{C}_{d,k}$ valós számot a következőképpen:

$$(1.2) \quad \tilde{C}_{d,k} := \sup_{1 < t < \infty} \left[t \log \left\{ \frac{2d}{(t-1) \left(\frac{t+1}{t-1} \right)^{k+1} - 1} \right\} \right].$$

Ekkor egy tetszőleges, az (1.1)-et kielégítő $p(z) = \sum_{j=0}^m a_j z^j$ ($\neq 0$) polinomra

$$(1.3) \quad \frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\theta})| d\theta - \log \left(\sum_{j=0}^m |a_j| \right) \geq \tilde{C}_{d,k}.$$

Megjegyzés. E tétel legfontosabb jellegzetessége az, hogy a $\tilde{C}_{d,k}$ alsó korlát független a $p(z)$ fokszámától.

Bizonyítás. Legyen $p(z) = \sum_{j=0}^m a_j z^j$ egy tetszőleges polinom ($\neq 0$), amely kielégíti (1.1)-et. Az általánosság megszorítása nélkül normalizálhatjuk $p(z)$ -t úgy, hogy $\sum_{j=0}^m |a_j| = 1$. Így (1.3) alapján elegendő bebizonyítani, hogy

$$(1.3') \quad \frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\theta})| d\theta \geq \tilde{C}_{d,k}.$$

Tetszőleges $0 < r < 1$ esetén a Cauchy formulából azt kapjuk, hogy

$$a_j = \frac{1}{2\pi} \int_0^{2\pi} \frac{p(re^{i\theta}) d\theta}{r^j e^{ij\theta}} \quad (j = 0, 1, \dots, m),$$

amiből világos, hogy abszolút értéket véve,

$$(1.4) \quad \sum_{j=0}^k |a_j| \leq \max_{|z|=r} \{|p(z)|\} \cdot \sum_{j=0}^k \frac{1}{r^j} = \max_{|z|=r} |p(z)| \cdot \left\{ \frac{\frac{1}{r^{k+1}} - 1}{\frac{1}{r} - 1} \right\}.$$

Mármost válasszunk egy tetszőleges z_0 -t úgy, hogy $|z_0| = r$ és $|p(z_0)| = \max_{|z|=r} |p(z)|$.

Vegyük észre, hogy $|p(z_0)| > 0$, hiszen $p(z) \neq 0$. Ezekután legyen $f(z)$ egy tetszőleges, a $|z| \leq 1$ körlemezen analitikus függvény úgy, hogy $f(0) \neq 0$, és legyenek $Z_\Delta(f)$ az $f(z)$ gyökei a $0 \leq |z| < 1$ nyílt körlemezen. Ekkor a *Jensen formula* (lásd Ahlfors [1, 207. oldal]) az

$$(1.5) \quad \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| d\theta = \log |f(0)| + \sum_{z_j \in Z_\Delta(f)} \log \left(\frac{1}{|z_j|} \right)$$

alakot ölti. Mivel itt az utolsó tag nemnegatív, klasszikus *Jensen egyenlőtlenség* az alábbi formában írható fel:

$$(1.5') \quad \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| d\theta \geq \log |f(0)|.$$

Speciálisan az adott $p(z)$ polinommal és a $w(z) := \frac{z+z_0}{1+\bar{z}_0 z}$ Möbius függvénnyel (amely a $|z| \leq 1$ körlemez konformisan képezi le a $|w| \leq 1$ körlemezre), legyen $f(z) := p(w(z)) = p\left(\frac{z+z_0}{1+\bar{z}_0 z}\right)$, ami analitikus a $|z| \leq 1$ tartományban, és $f(0)=p(z_0) \neq 0$. Ekkor az (1.5') alatti Jensen egyenlőtlenséget az $f(z)$ függvényre alkalmazva közvetlenül adódik, hogy

$$(1.6) \quad I := \frac{1}{2\pi} \int_0^{2\pi} \log \left| p \left(\frac{e^{i\theta} + z_0}{1 + \bar{z}_0 e^{i\theta}} \right) \right| d\theta \geq \log |p(z_0)|.$$

Innen az $e^{i\phi} := \left(\frac{e^{i\theta} + z_0}{1 + \bar{z}_0 e^{i\theta}} \right)$ változó-transzformációval és egy kis számolással kapjuk, hogy

$$(1.7) \quad \begin{aligned} I &= \frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\phi})| \cdot \frac{e^{i\phi}(1-r^2)}{\{-\bar{z}_0 e^{2i\phi} + (1+r^2)e^{i\phi} - z_0\}} d\phi \\ &= \frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\phi})| \cdot \frac{1-r^2}{|1-\bar{z}_0 e^{i\phi}|^2} d\phi. \end{aligned}$$

Itt az utóbbi integrandus második tényezőjére az alábbi *alsó korlátot* kapjuk:

$$(1.7') \quad \frac{1-r^2}{|1-\bar{z}_0 e^{i\phi}|^2} \geq \frac{1-r}{1+r} \quad (\text{bármely valós } \phi\text{-re}),$$

s mivel $|p(e^{i\phi})| = \left| \sum_{j=0}^m a_j e^{ij\phi} \right| \leq \sum_{j=0}^m |a_j| = 1$, továbbá ugyanezen integrálban az első tényező kielégíti a $\log |p(e^{i\phi})| \leq 0$ egyenlőtlenséget. Így, felhasználva az (1.7') egyenlőtlenséget, a következő *felső korlát* adódik I -re:

$$(1.8) \quad I \leq \frac{1}{2\pi} \left(\frac{1-r}{1+r} \right) \int_0^{2\pi} \log |p(e^{i\phi})| d\phi.$$

Kombinálva az (1.1) egyenlőtlenséget (a $\sum_{j=0}^m |a_j| = 1$ normalizálással) az (1.4), (1.6) és (1.8) relációkkal, azt kapjuk, hogy

$$\frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\phi})| d\phi \geq \left(\frac{1+r}{1-r} \right) \log \left[\frac{d\left(\frac{1}{r} - 1\right)}{\frac{1}{r^{k+1}} - 1} \right] \quad (0 < r < 1).$$

Itt a $t := \frac{1+r}{1-r}$ változó-transzformációt alkalmazva (lásd (1.2)),

$$(1.9) \quad \frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\phi})| d\phi \geq t \log \left\{ \frac{2d}{(t-1)\left(\frac{t+1}{t-1}\right)^{k+1} - 1} \right\} =: f_{d,k}(t),$$

bármely $1 < t < \infty$ esetén. Az $f_{d,k}(t)$ függvény nyilván valós értékű és folytonos az $(1, +\infty)$ intervallumon, és könnyen belátható, hogy

$$(1.10) \quad \tilde{C}_{d,k} := \sup_{1 < t < \infty} f_{d,k}(t)$$

végül bármely $0 < d < 1$ és nemnegatív egész k esetén. Így, mivel az (1.9) egyenlőtlenség teljesül bármely $1 < t < \infty$ mellett, ezért igaz az (1.10) alatt definiált $\tilde{C}_{d,k}$ számra is, ami nem más, mint a kívánt (1.3') egyenlőtlenség. \square

Az 1. Tétel egy érdekes következményeként tekintsük a $k = 0$ speciális esetet. Az (1.9) alapján

$$f_{d,0}(t) = t \log d \quad (1 < t < \infty),$$

úgyhogy

$$\tilde{C}_{d,0} := \sup_{1 < t < \infty} f_{d,0}(t) = \log d.$$

Vagyis bármely olyan $p(z) = \sum_{j=0}^m a_j z^j (\neq 0)$ polinomra, amelyre $|a_0| = d \sum_{j=0}^m |a_j|$, az (1.3) reláció az

$$\frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\theta})| d\theta \geq \log \left\{ d \cdot \sum_{j=0}^m |a_j| \right\} = \log |a_0|$$

relációra redukálódik, ami nem más, mint az (1.5') alatti Jensen egyenlőtlenség. Tehát (1.3) úgy is felfogható, mint a Jensen egyenlőtlenség egy általánosítása!

Mint azt Beazamy [3] megjegyzi, az $f_{d,k}(t)$ kiszámítása $t = 2$ esetén (lásd (1.9)) az alábbi alsó korlátot adja $\tilde{C}_{d,k}$ -ra:

$$(1.11) \quad \tilde{C}_{d,k} \geq 2 \log \left\{ \frac{2d}{3^{k+1} - 1} \right\}$$

bármely $0 < d < 1$ és nemnegatív egész k esetén. Mint az [3]-ban említést nyer, az 1. Tétel bizonyítása csak azt a (gyengébb) hipotézist használja, hogy $p(z) = \sum_{j=0}^m a_j z^j (\neq 0)$ kielégíti a

$$(1.12) \quad \sum_{j=0}^k |a_j| \geq d \quad \text{és} \quad \|p\|_{L^\infty(\Delta)} := \sup\{|p(z)| : |z| \leq 1\} \leq 1$$

egyenlőtlenségeket, úgyhogy az 1. Tétel állításai igazak ezen feltételezések mellett is.

Az 1. Tétel (1.3) egyenlőtlensége igen fontos, és felveti azt a nyilvánvaló kérdést, hogy határozzuk meg azt a legnagyobb valós $C_{d,k}$ számot, amelyre (1.3) teljesül bármely, az (1.1)-et kielégítő $p(z) (\neq 0)$ polinomra. Így, ha bevezetjük a

$$(1.13) \quad J(p) := \frac{1}{2\pi} \int_0^{2\pi} \log |p(e^{i\theta})| d\theta - \log \left(\sum_{j=0}^m |a_j| \right)$$

funkcionált, akkor

$$(1.14) \quad C_{d,k} := \inf \left\{ J(p) : p(z) = \sum_{j=0}^m a_j z^j (\neq 0) \text{ kielégíti (1.1)-et} \right\}.$$

Egy másik nyilvánvaló kérdés az *extremális polinomok*, vagyis azon (1.1)-et kielégítő $p(z)$ -k jellemzése, amelyekre $C_{d,k} = J(p)$. Ez a kérdés *nyitott* e könyv írásának pillanatában.

Igen szemléletes a $d = 1/2$ speciális eset vizsgálata. Minden egyes pozitív egész k esetén a speciális $(1+z)^{2k+1}$ polinom kielégíti (1.1)-et (egyenlőségjellel) $d = 1/2$ esetén. Továbbá, mivel $(1+z)^{2k+1}$ -nek nincs gyöke $|z| < 1$ -ben, ezért az (1.5) Jensen formula (1.13) és (1.14) segítségével azt adja, hogy

$$(1.15) \quad J((1+z)^{2k+1}) = -(2k+1) \log 2 \geq C_{1/2,k} \quad (k = 1, 2, \dots).$$

Mint azt a következő paragrafusban látni fogjuk, az (1.15) alatti egyenlőtlenség valójában *éles* a polinomok egy széles osztályára, minden egyes pozitív egész k esetén. Mindazonáltal van egy nagy *hézag* a $C_{1/2,k}$ -ra vonatkozó (1.15)-beli felső, és a $\tilde{C}_{1/2,k}$ -ra vonatkozó 1. Tételbeli alsó korlát között. Nevezetesen, a $k = 1$ esetben a $C_{1/2,1}$ -re vonatkozó (1.15) alatti felső korlát és a $\tilde{C}_{1/2,1}$ alsó korlát (amelyet (1.10)-ből numerikusan határozhatunk meg) a következők:

$$(1.16) \quad -2,07944 \dots \geq C_{1/2,1} \geq \tilde{C}_{1/2,1} = -3,69263 \dots$$

6.2. A Hurwitz polinomokkal kapcsolatos eredmények

E részben ismertetjük Rügler, Trimble és Varga [12] eredményeit a Jensen egyenlőtlenség Beauzamy–Enflo féle kiterjesztésével kapcsolatban, amelyek a speciális *H Hurwitz polinomokra* (lásd Marden [10, 181. oldal]) vonatkoznak. Ezeket céljainknak megfelelően a következőképpen normalizáljuk:

$$(2.1) \quad H := \left\{ \begin{array}{l} f(z) = \sum_{j=0}^m a_j z^j : m \text{ egy tetszőleges nemnegatív egész,} \\ f(0) = 1, \text{ az } a_j \text{ együtthatók mind valósak} \\ (j = 0, 1, \dots, m), \text{ és az } f(z) \text{ összes gyöke} \\ \text{a } \operatorname{Re} z < 0 \text{ tartományban van.} \end{array} \right.$$

Az (1.14) analógiájára, tetszőleges $0 < d < 1$ és nemnegatív egész k esetén legyen

$$(2.2) \quad C_{d,k}^H := \inf \{ J(f) : f(z) \in H \text{ és } f(z) \text{ kielégíti (1.1)-et} \},$$

ahol általában, az (1.13) és az (1.5) Jensen formula szerint

$$(2.3) \quad J(f) := \log \left\{ \frac{|a_0|}{\sum_{j=0}^m |a_j| \cdot \prod_{z_j \in \mathbb{Z}_\Delta(f)} |z_j|} \right\}.$$

Mivel H részhalmaza az *összes*, (1.1)-et kielégítő polinomnak, azért az (1.14)-ből és (2.2)-ből világos, hogy

$$(2.4) \quad C_{d,k}^H \geq C_{d,k}.$$

Másrészt a H elemeinek speciális szerkezete miatt, a [12]-beli eredményekből kiderül, hogy (i) az összes $C_{d,k}^H$ konstans *expliciten meghatározható*, és hogy (ii) a H -beli összes *extremális elem*, vagyis azok a $q(z) \in H$ polinomok, melyekre

$$(2.5) \quad J(q) = C_{d,k}^H,$$

úgyszintén meghatározhatók.

E célból először is, mint (2.1) közvetlen következményét megjegyezzük, hogy ha $f(z) = \sum_{j=0}^m a_j z^j \in H$, és ha ∂f jelöli az $f(z)$ *pontos fokát*, akkor

$$(2.6) \quad \begin{cases} \text{(i)} & a_0 = 1, \text{ és } a_j > 0 \text{ minden } j = 0, 1, \dots, \partial f \text{ esetén;} \\ \text{(ii)} & \text{ha } f(\rho) = 0 \text{ ahol } \rho < 0, \text{ akkor } f(z) / \left(1 - \frac{z}{\rho}\right) \in H; \\ \text{(iii)} & \text{ha } \rho < 0, \text{ akkor } f(z) \left(1 - \frac{z}{\rho}\right) \in H; \\ \text{(iv)} & \text{ha } f(\rho) = 0 \text{ ahol } \rho \text{ nem valós,} \\ & \text{akkor } f(z) / [(1 - \frac{z}{\rho})(1 - \bar{\rho}z)] \in H. \end{cases}$$

Mármost, ha $f(z) = \sum_{j=0}^m a_j z^j \in H$, akkor (2.6i) maga után vonja, hogy $f(1) = \sum_{j=0}^m a_j = \sum_{j=0}^m |a_j|$, és mivel $a_0 = 1$ a (2.6i) alapján, a (2.3) a

$$(2.7) \quad J(f) = -\log \left(f(1) \cdot \prod_{z_j \in Z_{\Delta}(f)} |z_j| \right) \quad (f(z) \in H)$$

alakra redukálódik. Célszerű lesz definiálni az alábbi számokat *bármely* $f(z) = \sum_{j=0}^m a_j z^j (\neq 0)$ esetén:

$$(2.8) \quad \delta_k(f) := \sum_{j=0}^k |a_j| / \sum_{j=0}^m |a_j| \quad (k = 0, 1, \dots, m).$$

Jegyezzük meg, hogy ha $f(z)$ k -adfokú koncentrációja d , akkor (lásd (1.1))

$$(2.9) \quad \delta_k(f) \geq d.$$

Az alábbi három, [12]-ből vett lemma szükséges lesz ezen rész fő eredményének az 5. Tételnek bizonyításához. Kezdjük a 2. Lemmával. Ennek lényege, vagyis a (2.10) reláció, megtalálható Mahler [8]-ban. Megjegyezzük, hogy a (2.10)-beli egyenlőség tárgyalása megtalálható [12]-ben. (Hasonló eredmények vannak Mahler [9] dolgozatában.)

2. LEMMA ([12] és Mahler[8]). Tekintsünk egy tetszőleges $f(z) = \sum_{j=N}^m a_j z^j$ komplex polinomot, ahol $a_N a_m \neq 0$. Ekkor (lásd (1.13)),

$$(2.10) \quad J(f) \geq -m \log 2,$$

és itt egyenlőség akkor és csak akkor fordul elő, ha $N = 0$ és $f(z) = \gamma(e^{i\psi} + z)^m$, ahol $\gamma \neq 0$ és ψ valós.

Bizonyítás. A $J(f)$ funkcionálra vonatkozó (2.3) általános formula alapján könnyű belátni, hogy tetszőleges olyan $g(z)$ és $h(z)$ polinomokra, amelyekre $g(0) \cdot h(0) \neq 0$ teljesül, igaz, hogy

$$J(gh) \geq J(g) + J(h).$$

A továbbiakban az általánosság megszorítása nélkül tekintsünk egy tetszőleges

$$f(z) = \sum_{j=0}^m a_j z^j = \prod_{j=1}^m (z + z_j) \text{ alakú komplex polinomot, ahol } |z_j| > 0 \ (j=1, 2, \dots, m).$$

Alkalmazva a fenti, induktíve kiterjesztett egyenlőtlenséget $f(z)$ -re, adódik, hogy

$$(2.10') \quad J(f) \geq \sum_{j=1}^m J(z + z_j).$$

Mármost, ha $|z_j| \geq 1$, akkor a (2.3) szerint $J(z + z_j) = \log(|z_j|/(1 + |z_j|)) \geq -\log 2$, ahol az utóbbi egyenlőtlenség következik az $u/(1+u)$ függvény $u \geq 1$ esetén fennálló monotonicitásából. Ha $0 < |z_j| < 1$, a (2.3) és egy hasonló okoskodás adja, hogy $J(z + z_j) = \log(1/(1 + |z_j|)) \geq -\log 2$. Alkalmazva ezen egyenlőtlenségeket a (2.10')-re

$$J(f) \geq -m \log 2$$

adódik, ami nem más, mint a kívánt (2.10) egyenlőtlenség. A (2.10)-ben szereplő szigorú egyenlőtlenség tárgyalása megtalálható [12]-ben. \square

3. LEMMA ([12]). Legyen k egy pozitív egész, és tekintsünk egy olyan $f(z) \in H$ függvényt, amelyre $\partial f := m \geq k + 1$. Tegyük fel, hogy z_1 és z_2 az $f(z)$ két (nem szükségképpen különböző) gyöke. Ekkor, kivéve ha z_1 és z_2 mindkettő valós úgy, hogy $z_1 = -1$ és $z_2 \leq -1$ (vagy fordítva), létezik egy $h(z) \in H$ úgy, hogy $\partial h = m$, és

$$(2.11) \quad J(f) > J(h)$$

továbbá (lásd (2.8))

$$(2.12) \quad \delta_k(h) > \delta_k(f).$$

Bizonyítás. Először is tegyük fel, hogy $\operatorname{Im} z_1$ and $\operatorname{Im} z_2$ közül legalább az egyik nem zérus, mondjuk $\operatorname{Im} z_1 \neq 0$. Az $f(z) \in H$ feltevés maga után vonja (lásd (2.1)), hogy \bar{z}_1 is egy gyöke az $f(z)$ -nek, és legyen $z_2 := \bar{z}_1$. Defináljuk a $g(z)$ és $h(z)$ polinomokat a

$$(2.13) \quad g(z) := \frac{f(z)}{\left(1 - \frac{z}{z_1}\right) \left(1 - \frac{z}{\bar{z}_1}\right)} \quad \text{és} \quad h(z) := \left(1 + \frac{z}{\rho}\right)^2 g(z)$$

formulákkal, ahol $\rho > 1$. A (2.6) alapján világos, hogy $g(z)$ és $h(z)$ mindketten elemei a H -nak.

Mivel $\partial f = m$, a (2.13) alapján írhatjuk, hogy $g(z) = \sum_{j=0}^{m-2} b_j z^j$, és mivel $g(z) \in H$ (2.6iv) miatt és $\partial g = m - 2$, ezért (lásd (2.6i)) $b_j > 0$ ($j = 0, 1, \dots, m - 2$). Jegyezzük meg, hogy az $m \geq k + 1$ feltevés maga után vonja, hogy $b_{k-1} > 0$. Ezekután egy rövid számolással kapjuk, hogy

$$\delta_k(f) = \left\{ \sum_{j=0}^{k-1} b_j + \left[(|z_1|^2 b_k - b_{k-1}) / |1 - z_1|^2 \right] \right\} / g(1),$$

és

$$\delta_k(h) = \left\{ \sum_{j=0}^{k-1} b_j + \left[(\rho^2 b_k - b_{k-1}) / (1 + \rho)^2 \right] \right\} / g(1).$$

Így $\delta_k(h) > \delta_k(f)$ akkor és csak akkor, ha

$$(2.14) \quad b_{k-1} \left[\frac{1}{|1 - z_1|^2} - \frac{1}{(1 + \rho)^2} \right] > b_k \left[\frac{|z_1|^2}{|1 - z_1|^2} - \frac{\rho^2}{(1 + \rho)^2} \right].$$

Az $f(z)$ 1-nél kisebb abszolút értékű gyökeinek halmazát ismét $Z_\Delta(f)$ -fel jelölve, legyen $Z' := Z_\Delta(f) \setminus \{z_1, \bar{z}_1\}$. Ekkor (2.7) és (2.13) alapján

$$J(f) = \log \left[\frac{\max\{|z_1|^2; 1\}}{g(1)|1 - z_1|^2 \prod_{\zeta \in Z'} |\zeta|} \right],$$

és

$$J(h) = \log \left[\frac{\rho^2}{g(1)(1 + \rho)^2 \prod_{\zeta \in Z'} |\zeta|} \right].$$

Ilymódon $J(f) > J(h)$ akkor és csak akkor, ha

$$(2.15) \quad \frac{\max\{|z_1|; 1\}}{|1 - z_1|} > \frac{\rho}{1 + \rho}.$$

Ha $|z_1| < 1$, akkor $1/2 < 1/|1 - z_1| < 1$, mivel $\operatorname{Re} z_1 < 0$. Ezért létezik egy $\rho > 1$ úgy, hogy

$$(2.16) \quad \frac{1}{|1 - z_1|} > \frac{\rho}{1 + \rho} > \frac{|z_1|}{|1 - z_1|}.$$

Itt az első egyenlőtlenség azt mutatja, hogy (2.15) teljesül, és mivel $\rho > 1$, azért $1/|1 - z_1| > 1/(1 + \rho)$. Így (2.14)-ben a b_{k-1} szorzója *pozitív*. Másrészt a (2.16)-beli utolsó egyenlőtlenség szerint ugyanitt a b_k szorzója *negatív*. Mint már láttuk, $b_{k-1} > 0$, és (2.6i) alapján következik, hogy $b_k \geq 0$. Következésképpen (2.14) úgyszintén teljesül, és ezzel a 3. Lemmában a (2.11) és (2.12) ebben az esetben bizonyítást nyert.

Ha $|z_1| \geq 1$, akkor $1/2 < (|z_1|/|1 - z_1|) < 1$, mivel $\operatorname{Re} z_1 < 0$ és $z_1 \neq -1$. Ezért létezik egy $\rho_1 > 1$ úgy, hogy

$$\frac{|z_1|}{|1 - z_1|} = \frac{\rho_1}{1 + \rho_1}.$$

Ilymódon $1 + 1/\rho_1 = |1 - 1/z_1| < 1 + 1/|z_1|$. Ebből következik, hogy $\rho_1 > |z_1|$, vagyis $1/|1 - z_1| > 1/(1 + \rho_1)$. Így a (2.14) jobb oldala zérus, ha $\rho = \rho_1$, és a bal oldal pozitív. A folytonosságból következik, hogy létezik egy ρ az $(1, \rho_1)$ intervallumban úgy, hogy (2.14) és (2.15) egyszerre teljesül. Következésképpen abból, hogy z_1 egy olyan gyöke $f(z)$ -nek, hogy $\operatorname{Im} z_1 \neq 0$, adódik egy olyan $h(z) \in H$ létezése, amelyre a 3. Lemma (2.11) és (2.12) állítása egyidejűleg teljesül.

Mármost legyen $\operatorname{Im} z_1 = \operatorname{Im} z_2 = 0$. Három esetet különböztetünk meg. Először tegyük fel, hogy z_1 és z_2 mindketten a $(-1, 0)$ nyílt intervallumban vannak, mondjuk $-1 < z_1 < 0$. Változtassuk meg $g(z)$ és $h(z)$ definícióját a következőképpen:

$$g(z) := \frac{f(z)}{\left(1 - \frac{z}{z_1}\right)} \quad \left(\text{és legyen } g(z) = \sum_{j=0}^{m-1} b_j z^j \right),$$

és

$$h(z) := (1 + z)g(z).$$

A (2.14) bizonyításához hasonló módon adódik, hogy

$$\delta_k(f) = \left[\sum_{j=0}^{k-1} b_j - z_1 b_k / (1 - z_1) \right] / g(1),$$

és

$$\delta_k(h) = \left[\sum_{j=0}^{k-1} b_j + b_k/2 \right] / g(1).$$

Ebben az esetben $b_k > 0$ is teljesül. Így $\delta_k(h) > \delta_k(f)$ akkor és csak akkor, ha $1/2 > -z_1/(1-z_1)$, ami mindig igaz, hacsak $z_1 \in (-1, 0)$. Legyen $Z' := Z_\Delta(f) \setminus \{z_1\}$. A (2.7)-ből, valamint a $g(z)$ és $h(z)$ új definíciójából kapjuk, hogy

$$J(f) = \log \left[\frac{1}{g(1)(1-z_1) \prod_{\zeta \in Z'} |\zeta|} \right],$$

és

$$J(h) = \log \left[\frac{1}{2g(1) \prod_{\zeta \in Z'} |\zeta|} \right].$$

Ezért $J(f) > J(h)$ akkor és csak akkor, ha $1/(1-z_1) > 1/2$, és ez utóbbi egyenlőtlenség nyilván igaz. Ezzel az első eset bizonyítását befejeztük.

Mármost tegyük fel, hogy $\text{Im } z_1 = \text{Im } z_2 = 0$, továbbá hogy z_1 és z_2 mindketten a $(-\infty, -1)$ intervallumban vannak. Ezenkívül feltesszük, hogy

$$(2.17) \quad 1 - z_1 - z_2 - z_1 z_2 \geq 0.$$

A $g(z)$ és $h(z)$ definíciója most így módosul:

$$g(z) := \frac{f(z)}{(1 - \frac{z}{z_1})(1 - \frac{z}{z_2})} \quad \left(\text{és legyen } g(z) = \sum_{j=0}^{m-2} b_j z^j \right),$$

és

$$h(z) := (1+z) \left(1 + \frac{z}{\rho} \right) g(z), \text{ ahol } \rho > 1.$$

Mint a (2.14) bizonyításánál, $\delta_k(h) > \delta_k(f)$ akkor és csak akkor, ha

$$(2.18) \quad \begin{aligned} & b_{k-1} \left[\frac{1}{(1-z_1)(1-z_2)} - \frac{1}{2(1+\rho)} \right] \\ & > b_k \left[\frac{z_1 z_2}{(1-z_1)(1-z_2)} - \frac{\rho}{2(1+\rho)} \right]. \end{aligned}$$

Mint a (2.15) bizonyításánál, $J(f) > J(h)$ akkor és csak akkor, ha

$$(2.19) \quad \frac{z_1 z_2}{(1-z_1)(1-z_2)} > \frac{\rho}{2(1+\rho)}.$$

Ha (2.17)-ben egyenlőség áll fenn, akkor (2.19) bal oldala egyenlő lesz $1/2$ -del, és (2.19) fennáll bármely $\rho \geq 1$ esetén. Továbbá, mivel (2.18)-ban a b_k szorzója pozitív és zérushoz tart, midőn $\rho \rightarrow \infty$, és mivel hasonlóképpen kapjuk, hogy $b_{k-1} > 0$, innen következik, hogy a (2.18)-beli egyenlőtlenség realizálható azáltal, hogy $\rho \geq 1$ értékét megfelelően nagyra választjuk.

Ezekután tegyük fel, hogy (2.17)-ben szigorú egyenlőtlenség teljesül. Ezt összevetve azzal a ténnyel, hogy z_1 és z_2 a $(-\infty, -1)$ intervallumban vannak, következik, hogy $1/4 < z_1 z_2 / [(1 - z_1)(1 - z_2)] < 1/2$. Következésképpen létezik egy $\rho_2 > 1$ úgy, hogy

$$\frac{z_1 z_2}{(1 - z_1)(1 - z_2)} = \frac{\rho_2}{2(1 + \rho)}.$$

Ebből következik, hogy

$$\frac{1}{2(1 + \rho_2)} = \frac{2 - (z_1 + 1)(z_2 + 1)}{2(1 - z_1)(1 - z_2)} < \frac{1}{(1 - z_1)(1 - z_2)}.$$

Így (2.18) jobb oldala zérus, ha $\rho = \rho_2$, és mivel csakúgy, mint korábban, $b_{k-1} > 0$, a (2.18) bal oldala pozitív, ha $\rho = \rho_2$. A folytonosságból következik, hogy létezik egy $\rho \in (1, \rho_2)$ úgy, hogy (2.18) és (2.19) egyszerre teljesül. Ebből ismét csak következik, hogy létezik egy $h(z) \in H$, amely egyszerre kielégíti a 3. Lemmában a (2.11)-et és (2.12)-t. Végül tegyük fel, hogy $\operatorname{Im} z_1 = \operatorname{Im} z_2 = 0$, z_1 és z_2 mindketten a $(-\infty, -1)$ intervallumban vannak, de (2.17) nem teljesül. A korábban definiált $g(z)$ -t, valamint a $h(z) := (1 + z/\rho)g(z)$ ($\rho > 1$) függvényt fogjuk használni. Ekkor $\delta_k(h) > \delta_k(f)$ akkor és csak akkor, ha

$$(2.20) \quad \frac{b_{k-1}}{(1 - z_1)(1 - z_2)} > b_k \left[\frac{z_1 z_2}{(1 - z_1)(1 - z_2)} - \frac{\rho}{1 + \rho} \right],$$

és $J(f) > J(h)$ akkor és csak akkor, ha

$$(2.21) \quad \frac{z_1 z_2}{(1 - z_1)(1 - z_2)} > \frac{\rho}{1 + \rho}.$$

Mivel (2.17) nem teljesül, azért $1/2 < z_1 z_2 / [(1 - z_1)(1 - z_2)] < 1$. Így létezik egy $\rho_3 > 1$ úgy, hogy

$$\frac{z_1 z_2}{(1 - z_1)(1 - z_2)} = \frac{\rho_3}{1 - \rho_3}.$$

A folytonosság miatt létezik egy $\rho \in (1, \rho_3)$ úgy, hogy a (2.20) és (2.21) egyszerre teljesül, s ezért létezik egy $h(z) \in H$, amely a 3. Lemma (2.11) és (2.12) állításait is kielégíti. \square

Érdemes megjegyezni, hogy a [12]-beli 3. Lemmát úgyszintén a MACSYMA programcsomagot használó komputer-tapasztalatok motiválták, ahol egy, az (1.1) koncentráció-korlátozást kielégítő $p(z) \in H$ polinom gyökeit páronként úgy mozgattuk, hogy az (1.13)-beli $J(p)$ funkcionál csökkent, miközben az (1.1) korlátozást változatlanul kielégítettük.

4. LEMMA ([12]). Tetszőleges $d \in (0, 1)$ és k nemnegatív egész esetén létezik egy egyértelmű pozitív egész n , amely d -től és k -tól függ úgy, hogy

$$(2.22) \quad \frac{1}{2^n} \sum_{j=0}^k \binom{n}{j} \leq d < \frac{1}{2^{n-1}} \sum_{j=0}^k \binom{n-1}{j}.$$

Továbbá, ha

$$(2.23) \quad \hat{\rho} := \frac{\binom{n-1}{k}}{\sum_{j=0}^k \binom{n-1}{j} - d 2^{n-1}} - 1,$$

akkor $\hat{\rho} > 1$.

Bizonyítás. Egy adott nemnegatív egész k esetén tekintsük a pozitív valós számok egy

$$(2.24) \quad \left\{ \frac{1}{2^\ell} \sum_{j=0}^k \binom{\ell}{j} \right\}_{\ell=k}^{\infty}$$

sorozatot, amely 1-gyel kezdődik. Azt állítjuk, hogy e sorozat szigorúan monoton csökken, és határértéke zérus. E célból legyen

$$(2.25) \quad a_\ell := \frac{1}{2^\ell} \sum_{j=0}^k \binom{\ell}{j} \quad (\ell = k, k+1, \dots).$$

Mivel

$$(2.26) \quad \binom{\ell+1}{j} = \binom{\ell}{j} + \binom{\ell}{j-1} \quad (j = 1, 2, \dots, \ell),$$

a (2.25)-ből következik, hogy

$$a_{\ell+1} = a_\ell - \frac{1}{2^{\ell+1}} \binom{\ell}{k} \quad (\ell = k, k+1, \dots),$$

vagyis a (2.24) elemei szigorúan csökkennek. Mármost a *központi határeloszlástétel* egy szép következményeképpen (lásd Patel és Read [11, 169–170. oldal]) adódik, hogy

$$(2.27) \quad \left| a_\ell - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(2k+1-\ell)/\sqrt{\ell}} e^{-t^2/2} dt \right| < \frac{0,28}{\sqrt{\ell}}$$

minden $\ell \geq \max\{k; 1\}$ esetén. Mivel k rögzített, (2.27) mutatja, hogy $a_\ell \rightarrow 0$ midőn $\ell \rightarrow +\infty$. (Egy egyszerűbb bizonyítás arra, hogy $a_\ell \rightarrow 0$, midőn $\ell \rightarrow \infty$ az, hogy (2.25)-ből közvetlenül következik: bármely $\ell \geq 2k$ esetén fennáll, hogy

$$0 < a_\ell < \frac{(k+1)\ell^k}{k!2^\ell},$$

és mivel k rögzített, e felső korlát nyilván zérushoz tart, midőn $\ell \rightarrow \infty$.) Így adott egy tetszőleges $d \in (0, 1)$, a (2.25) alatti a_ℓ szigorúan csökkenő volta maga után vonja egy egyértelmű pozitív egész n létezését úgy, hogy $n \geq k+1$, és (2.22) teljesül. A (2.22) és (2.26)-ból közvetlenül következik, hogy a (2.23) által definiált $\hat{\rho}$ -ra teljesül $\hat{\rho} \geq 1$. \square

Ezzel elérkeztünk ennek a pontnak a főeredményéhez.

5. TÉTEL ([12]). *Adott egy tetszőleges $d \in (0, 1)$ és egy tetszőleges pozitív egész k , legyen n és $\hat{\rho}$ (egyértelműen) definiálva a (2.22) és (2.23) által. Ekkor (lásd (2.2))*

$$(2.28) \quad C_{d,k}^H = \log \left[\frac{\hat{\rho}}{(\hat{\rho}+1)2^{n-1}} \right].$$

Továbbá legyen $Q_{n,\hat{\rho}}(z) := \left(1 + \frac{z}{\hat{\rho}}\right)(1+z)^{n-1}$, ami a H egy eleme. Ekkor az (1.1)-et kielégítő $f(z) \in H$ egy extrémális elem a H -ban (vagyis $J(f) = C_{d,k}^H$) akkor és csak akkor, ha $f(z) = Q_{n,\hat{\rho}}(z)$.

Bizonyítás. Adott egy $d \in (0, 1)$ és egy pozitív egész k , legyenek a pozitív egész n és valós $\hat{\rho} \geq 1$ a 4. Lemma (2.22) és (2.23) relációi által egyértelműen definiált mennyiségek. Ekkor a $Q_{n,\hat{\rho}}(z) := \left(1 + \frac{z}{\hat{\rho}}\right)(1+z)^{n-1}$ polinom a H halmazhoz tartozik, és belátható, hogy

$$(2.29) \quad \delta_k(Q_{n,\hat{\rho}}) = d \quad \text{és} \quad J(Q_{n,\hat{\rho}}) = \log \left[\frac{\hat{\rho}}{(\hat{\rho}+1)2^{n-1}} \right].$$

Definíció szerint (lásd (2.2)) a (2.29) maga után vonja, hogy

$$(2.30) \quad J(Q_{n,\hat{\rho}}) = \log \left[\frac{\hat{\rho}}{(\hat{\rho}+1)2^{n-1}} \right] \geq C_{d,k}^H.$$

Mármost, mivel $\log \left[\frac{\rho}{(\rho+1)2^{n-1}} \right]$ mint ρ függvénye szigorúan monoton növekszik, ha $\rho \geq 1$, és felülről korlátos a $-(n-1)\log 2$ korláttal, a 2. Lemmából következőleg nem kell H -beli, n -nél alacsonyabbfokú polinomokat tekinteni. Ezért az n (2.22) alatti definíciójából következik (lásd (2.24)), hogy $n \geq k+1$. Így a 3. Lemma alapján elegendő tekinteni egy $f(z) = \left(1 + \frac{z}{\rho}\right)(1+z)^{m-1}$ függvényt, ahol $m \geq n$ és $\rho \geq 1$.

Mivel ezen $f(z)$ -nek ki kell elégítenie (1.1)-et, megmutatható, hogy $m \leq n$, és ha $m = n$, akkor $\hat{\rho} \leq \rho$, ahol $\hat{\rho}$ a (2.23) alatt definiált mennyiség. Így csak azt az esetet kell tekinteni, amikor $m = n$ és $\hat{\rho} \leq \rho$. Egy, a (2.7)-re alapozott számítás mutatja, hogy

$$(2.31) \quad J \left(\left(1 + \frac{z}{\rho} \right) (1 + z)^{n-1} \right) = \log \left\{ \frac{\rho}{(1 + \rho)^{2^{n-1}}} \right\}.$$

De a (2.31) jobb oldalán álló mennyiség ismét csak szigorúan monoton növekedő, ha $\rho \geq \hat{\rho}$. Így a fenti mennyiség minimalizálásához a $\rho = \hat{\rho}$ választás szükséges, ami nemcsak azt adja, hogy

$$(2.32) \quad \inf \{ J(f) : f(z) \in H \text{ és } f(z) \text{ kielégíti (1.1)-et} \} = J(Q_{n,\hat{\rho}}) = C_{d,k}^H,$$

hanem azt is, hogy $Q_{n,\hat{\rho}}(Z)$ az az *egyértelmű* elem H -ban, amelyre $J(Q_{n,\hat{\rho}}) = C_{d,k}^H$. \square

E rész befejezéseként megemlítiük, hogy nagyszámú numerikus kísérlet arra a meggyőződésre vezetett bennünket, hogy (2.4)-ben bizonyára egyenlőség áll fenn, s ezért kimondjuk a következő sejtést:

$$(2.33) \quad \text{Sejtés (Rigler, Trimble és Varga [12]): } C_{d,k}^H \stackrel{?}{=} C_{d,k} \\ \text{minden } 0 < d < 1 \text{ és minden } k = 0, 1, \dots \text{ esetén.}$$

6.3. Általánosítások és megjegyzések

Annak a ténynek az ismeretében, hogy a 6.1 és 6.2 részek polinomokkal kapcsolatos eredményei a tekintett polinomok *fokszáma*it nem korlátozzák, nem meglepő, hogy ezen eredmények *változtatás nélkül* átvihetők bizonyos *zérus rendű* egész függvényekre (lásd Boas [5]). Mint azt [12]-ben megmutattuk, az 5. Tétel eredményei valóban igazak az alábbi halmaz elemeire:

$$\tilde{H} := \left\{ f(z) = \prod_{j=1}^{\omega} \left(1 - \frac{z}{z_j} \right) \text{ (ha } \omega \leq \infty \text{) : } z_k \in \{z_j\}_{j=1}^{\omega} \right.$$

$$\left. \text{maga után vonja, hogy } \operatorname{Re} z_k < 0 \text{ és } \bar{z}_k \in \{z_j\}_{j=1}^{\omega}, \text{ továbbá } \sum_{j=1}^{\omega} \frac{1}{|z_j|} < \infty \right\}.$$

Jegyezzük meg, hogy \tilde{H} magában foglalja a (2.1)-beli H -t, mint egy *valódi részhalmazt*.

Egy másik lehetséges kiterjesztés az, hogy túllépünk a zárt $|z| \leq 1$ egységgörön analitikus függvények halmazán. Legyen $f(z) = \sum_{j=0}^{\infty} a_j z^j$ analitikus $|z| < 1$ -ben, és legyen

$$(3.1) \quad M_{\infty}(r; f) := \max_{0 \leq \theta \leq 2\pi} |f(re^{i\theta})|, \text{ minden egyes } 0 \leq r < 1 \text{ esetén.}$$

Szokásos módon a H^{∞} Hardy-teret (lásd Duren [7, 2. oldal]) úgy definiáljuk, mint

$$(3.2) \quad H^{\infty} := \{g(z) : g(z) \text{ analitikus } |z| < 1\text{-ben és} \\ M_{\infty}(r; g) \text{ korlátos, midőn } r \rightarrow 1-\}.$$

Tegyük fel, hogy egy, a $|z| < 1$ körben analitikus $f(z) = \sum_{j=0}^{\infty} a_j z^j$ függvény esetén vizsgáljuk az (1.1) fogalom megfelelő voltát. Legyen ezen $f(z)$ k -adfokú koncentrációja d . Hogy ez jelentéssel bírjon, ahhoz szükséges $\sum_{j=0}^{\infty} |a_j| < \infty$ fennállása, s így (lásd (3.1))

$$M_{\infty}(r; f) \leq \sum_{j=0}^{\infty} |a_j| r^j \leq \sum_{j=0}^{\infty} |a_j| < \infty;$$

amiből (3.2) alapján $f(z) \in H^{\infty}$. Megmutatható (lásd [7, 17. oldal]), hogy bármely $f(z) \in H^{\infty}$ kiterjeszthető a $|z| = 1$ határvonalra egy, a $[0, 2\pi]$ intervallumon definiált $\hat{f}(e^{i\theta})$ függvény segítségével, amelyre

$$(3.3) \quad \begin{cases} \hat{f}(e^{i\theta}) = \lim_{r \rightarrow 1-} f(re^{i\theta}) \text{ majdnem mindenütt } [0, 2\pi]\text{-ben;} \\ \hat{f}(e^{i\theta}) \in L^{\infty}[0, 2\pi]; \\ \text{ha } f(z) \neq 0, \text{ akkor } \log |\hat{f}(e^{i\theta})| \in L^1[0, 2\pi]. \end{cases}$$

Így egy tetszőleges $f(z) = \sum_{j=0}^{\infty} a_j z^j$ ($\neq 0$) esetén, amely analitikus a $|z| < 1$ körben és $\sum_{j=0}^{\infty} |a_j| < \infty$, fennáll, hogy $f(z) \in H^{\infty}$ és teljesül az (1.13) analagonja, vagyis

$$(3.4) \quad \hat{J}(f) := \frac{1}{2\pi} \int_0^{2\pi} \log |\hat{f}(e^{i\theta})| d\theta - \log \left(\sum_{j=0}^{\infty} |a_j| \right)$$

jól definiált és véges. Ennek megfelelően arra a problémára jutunk, hogy minimalizáljuk a (3.4)-beli $\hat{J}(f)$ funkcionált az összes olyan $f(z) \in H^{\infty}$ esetén, mely kielégíti a $\sum_{j=0}^k |a_j| \geq d \sum_{j=0}^{\infty} |a_j|$ relációt. Abban az esetben, amikor $0 < d < 1$ és

$k = 0$, ez a minimalizálás akkor és csak akkor következik be (lásd [12, 1. Tétel]), ha $|a_0| = d \sum_{j=0}^{\infty} |a_j|$ és $f(z)$ megegyezik saját *külső függvényével* (lásd Rudin [13, 338. oldal]).

Az a fogalom, hogy a $p(z) = \sum_{j=0}^m a_j z^j$ polinom k -adfokú koncentrációja d , vagyis hogy (lásd (1.1))

$$(3.5) \quad \sum_{j=0}^k |a_j| \geq d \sum_{j=0}^m |a_j|,$$

könnyen általánosítható, mivel a fenti egyenlőtlenségben csak az $(a_0, a_1, \dots, a_m)^T$ együtthatókból alkotott vektor, valamint e vektorok k -*csonkítottjának*, vagyis az $(a_0, a_1, \dots, a_k, 0, \dots, 0)^T$ vektornak az ℓ_1 -normája szerepel. Ehelyett nyilván tekinthetünk a (3.5)-ben ℓ_q -normákat, vagyis a

$$(3.6) \quad \left(\sum_{j=0}^k |a_j|^q \right)^{1/q} \geq d \left(\sum_{j=0}^m |a_j|^q \right)^{1/q} \quad (1 \leq q \leq \infty),$$

normát a megfelelő minimalizálási problémákkal együtt. Vagy akár keverhetjük is a normákat (lásd (1.12)):

$$(3.7) \quad \left(\sum_{j=0}^k |a_j|^q \right)^{1/q} \geq d \|p(z)\|_{L^\infty(\Delta)} \quad (1 \leq q \leq \infty),$$

más, megfelelő minimalizációs problémákra. Ilyen típusú általánosításokat vizsgált Beuzamy [4] és Bonvalot [6].

Adott $0 < d < 1$ és adott nemnegatív egész k esetén jegyezzük meg, hogy léteznek olyan $p(z)$ polinomok, amelyek *nincsenek* H -ban, és amelyek kielégítik mind (1.1)-et, mind a

$$(3.8) \quad J(p) = C_{d,k}^H$$

relációt. Erre egy egyszerű példa — csakúgy, mint az 5. Tételben — a

$$(3.9) \quad Q_{n,1}(z) := (1+z)^3 = 1 + 3z + 3z^2 + z^3,$$

speciális harmadfokú polinom, ami a \mathcal{H} halmazban van, és ami kielégíti (1.1)-et (egyenlőséggel), hacsak $d = 1/2$ és $k = 1$. Az 5. Tétel (2.28) relációjából következik, hogy

$$(3.10) \quad J(Q_{n,1}) = -3 \log 2 = C_{1/2,1}^H.$$

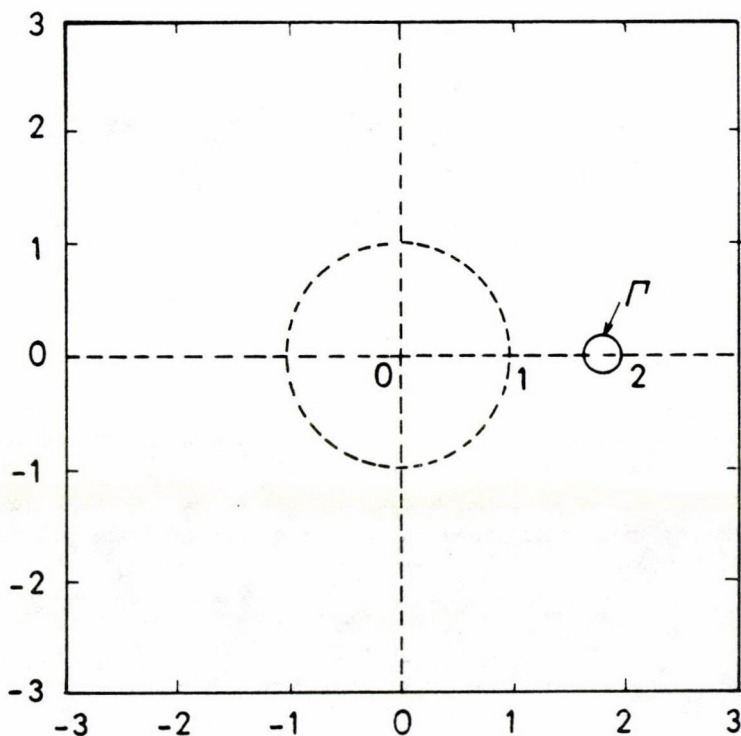
Másrészt tekintsük a

$$(3.11) \quad p_5(z) := (1+z)^3(4-z^2) = 4 + 12z + 11z^2 + z^3 - 3z^4 - z^5,$$

ötödfokú polinomot, amely nincs a (2.1)-beli H -ban, mivel $p_5(z)$ -nek van egy gyöke a $\operatorname{Re} z > 0$ tartományban. De $p_5(z)$ úgyszintén kielégíti (1.1)-et (egyenlőségjellel), hacsak $d = 1/2$ és $k = 1$, s ezenkívül (2.3) miatt $p_5(z)$ -re teljesül a

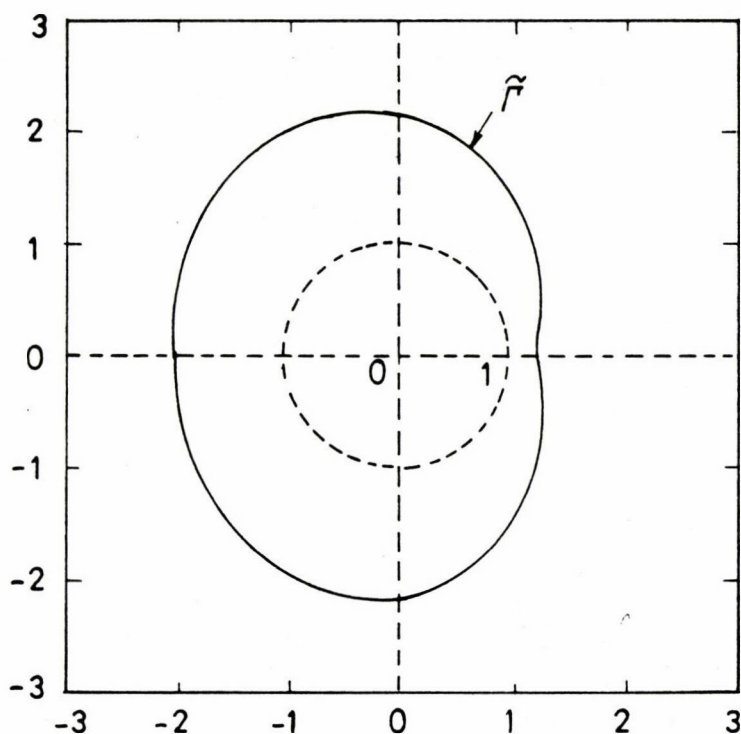
$$(3.12) \quad J(p_5) = -3 \log 2 = C_{1/2,1}^H$$

reláció is. Természetesen ebből következik, hogy ha a (2.23) sejtés igaz, vagyis $C_{d,k}^H = C_{d,k}$, akkor a $J(p) = C_{d,k}$ relációt kielégítő extrémális $p(z)$ polinomok nem szükségképpen egyértelműek, éles ellentétben az 5. Tétel eredményeivel!



6.1. ábra: A Γ görbe.

Végül annak illusztrálására, hogy a MACSYMA-val végzett tudományos számítások hogyan alkalmazhatók e fejezetben, bemutatjuk a 6.1 és 6.2 ábrákat, a következő magyarázattal. Vegyük a (3.11) alatti $p_5(z)$ polinomot, és változtassuk


 6.2. ábra: A $\tilde{\Gamma}$ görbe.

a $+2$ gyökét úgy, hogy közben kielégítjük az (1.1) korlátozást (egyenlőséggel) a $d = 1/2$ és $k = 1$ értékkel. Másszóval tekintsük a

$$(3.13) \quad P_{\mu}(z) := \{(1+z)^3(2+z)\} [\mu - z]$$

polinomot, ahol μ egy komplex paraméter a $+2$ kezdeti értékkel, és ahol $P_{\mu}(z)$ teljesíti azt a megkötést (lásd (2.8)), hogy $\delta_1(P_{\mu}) = 1/2$. Ekkor megmutatható, hogy létezik egy egyszerű zárt Γ görbe a komplex síkon, amely áthalad a $\mu = 2$ ponton, és amely a következőképpen van definiálva:

$$(3.14) \quad \Gamma := \{\mu \in \mathbb{C} : \delta_1(P_{\mu}) = 1/2\}.$$

Mivel (3.13)-ban a $[\mu - z]$ szorzója, $h(z) := \{(1+z)^3(2+z)\}$, kielégíti a $\delta_1(h)3/8 < 1/2$ relációt, megállapíthatjuk, hogy a Γ görbe nem terjeszthető ki a végtelenig.

A 6.1. ábrán bemutatjuk a Γ görbét, amelyről belátható, hogy

$$(3.15) \quad \min\{J(P_{\mu}) : \mu \in \Gamma\} = J(P_2) = -3 \log 2;$$

pontosabban a $J(P_{\mu})$ minimális értéke a Γ görbén csak a $\mu = 2$ értéknél vétetik fel. Így a Γ görbén végighaladva generálunk egy $P_{\mu}(z)$ polinomsorozatot, amelynek

minden tagja kielégíti a $\delta_1(P_\mu) = 1/2$ relációt, de a (3.12)-höz képest nincs javulás a megfelelő J funkcionál minimalizálásában e polinomsorozaton.

Vehetjük megint a (3.11) alatti $p_5(z)$ polinomot és változtassuk meg az egyszerűs -2 gyökét, miközben kielégítjük a $\delta_1(\tilde{P}_\mu) = 1/2$ korlátozást, ahol

$$\tilde{P}_\mu(z) := \{(1+z)^3(2-z)\}[-\tilde{\mu}+z],$$

és ahol $\tilde{\mu}$ egy komplex paraméter a -2 kezdeti értékkel. Ezen eljárás a

$$\tilde{\Gamma} := \{\tilde{\mu} \in \mathbb{C} : \delta_1(\tilde{P}_\mu) = 1/2\}$$

egyszerű zárt görbét szolgáltatja, amely áthalad a -2 ponton. Ezt szemléltettük a 6.2 ábrán. Csakúgy, mint (3.15)-nél, itt is

$$(3.16) \quad \min \{J(\tilde{P}_\mu) : \tilde{\mu} \in \tilde{\Gamma}\} = J(\tilde{P}_{-2}) = -3 \log 2.$$

IRODALOM

- [1] L.V. AHLFORS, *Complex Analysis*, harmadik kiadás (McGraw-Hill Book Co., New York, 1979).
- [2] B. BEAUZAMY és P. ENFLO, „Estimations de produits de polynômes”, *J. Number Theory* 21 (1985), 390–412.
- [3] B. BEAUZAMY, „Jensen's inequality for polynomials with concentration at low degrees”, *Number. Math.* 49 (1986), 221–225.
- [4] B. BEAUZAMY, „A minimization problem connected with the generalized Jensen's inequality”, *J. Math. Anal. Appl.* 145 (1990), 137–144.
- [5] R.P. BOAS, *Entire Functions* (Academic Press, Inc., New York, 1954).
- [6] L. BONVALOT, „Moyenne géométrique des fonctions des espaces de Hardy et polynômes concentrés aux bas degrés”, *Thèse de Troisième Cycle, Université de Paris* 7 (1986).
- [7] P.L. DURIN, *Theory of H^p Spaces* (Academic Press, Inc., New York, 1970).
- [8] K. MAHLER, „An application of Jensen's formula to polynomials”, *Mathematika* 7 (1960), 98–100.
- [9] K. MAHLER, „On two extremum properties of polynomials”, *III. J. Math.* 7 (1963), 618–701.
- [10] M. MARDEN, *Geometry of Polynomials*, Mathematical Surveys Number 3 (American Mathematical Society, Providence, R.I., 1966).
- [11] J.K. PATEL és C.B. READ, *Handbook of the Normal Distribution* (Marcel Dekker, New York, 1982).
- [12] A.K. RIGLER, S.Y. TRIMBLE és R.S. VARGA, „Sharp lower bounds for a generalized Jensen inequality”, *Rocky Mountain J. Math.* 19 (1989), 353–373.
- [13] W. RUDIN, *Real and Complex Analysis* (McGraw-Hill Book Co., New York, 1966).

RICHARD S. VARGA
INSTITUTE FOR COMPUTATIONAL MATHEMATICS
KENT STATE UNIVERSITY
KENT, OH 44242

Készült az ELTE Szokszorosítóüzemében
350 példányban
Felelős kiadó: Dr. Kiss Ádám
Felelős vezető: Arató Tamás
ELTE 94115

Fordította: Szabados József

Alkalmazott Matematikai Lapok 16 (1992)

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

ÚTMUTATÁS A SZERZŐKNEK

Az Alkalmazott Matematikai Lapok csak magyar nyelvű dolgozatokat közöl. A kéziratok gépelését, olyan formában kérjük, hogy minden gépelt oldal 25, egyenként átlag 50 betűhelyes sort tartalmazzon. A közlésre szánt dolgozatokat három példányban kell beküldeni. Előnyben részesülnek a TEX-ben elkészített dolgozatok. Ezeket két kinyomtatott példány kíséretében diszketten kérjük beadni.

A kéziratok szerkezeti felépítésének, a következő követelményeket kell kielégíteni. A fejlécnek tartalmaznia kell a dolgozat címét, a szerző teljes nevét, valamint annak a városnak a nevét, ahol a szerző dolgozik. A fejléc után egy, képletet nem tartalmazó, legfeljebb 200 szóból álló kivonatot kell minden esetben megadni. A dolgozatot címmel ellátott szakaszokra kell bontani, és az egyes szakaszokat arab sorszámmal kell ellátni. Az esetleges bevezetésnek mindig az első szakaszt kell alkotnia. Az irodalomjegyzék mindig az utolsó szakasz kell, hogy legyen, és azt nem kell sorszámmal ellátni. Az irodalomjegyzék után, a kézirat befejezésekképpen fel kell tüntetni a szerző teljes nevét és a munkahelye (illetve lakása) pontos postai címét. A dolgozatban előforduló képleteket szakaszonként újrakezdődően, a képlet előtt két zárójel közé írt kettős számozással kell azonosítani. Természetesen nem szükséges minden képletet számozással ellátni. Az esetleges definíciókat és tételeket (segéd tételeket és lemmákat) ugyancsak szakaszonként újrakezdődő, kettős számozással kell ellátni. Kérjük a szerzőket, hogy ezeket, valamint a tételek bizonyítását a szövegben kellő módon emeljék ki. Minden dolgozathoz csatolni kell egy angol, német, francia vagy orosz nyelvű, külön oldalra gépelt összefoglalót. Amennyiben lehetséges, kérjük a nyomtatás számára különösen nehézkes matematikai jelölések használatának az elkerülését.

A dolgozatok ábráit és az esetileges lábjegyzeteket a dolgozat végén, különálló lapokon kérjük beküldeni. Mind az ábrákat mind a lábjegyzeteket a dolgozat szakaszokra bontásától független, folytatódóan arab számozással kell ellátni. Az ábrák elhelyezését a dolgozat megfelelő helyén, széljegyzetként felüntetett, ábraszámokkal sorszámozott sorszámmal kell megadni. A lábjegyzeteket a dolgozaton belül az azonosító számok felső indexekénti használatával lehet hivatkozni.

A irodalmi hivatkozások formája a következő. Minden hivatkozást fel kell sorolni a dolgozat végén található irodalomjegyzékben, a szerzők, illetve társszerzők esetén az első szerző neve szerint alfabetikus sorrendben úgy, hogy a című betűs szerzők nevét a Mathematical Reviews átirási szabályai szerint latin betűsre kell írni. A foyóvatban megjelent cikkekre [1], a könyvekre [5], a kötetben megjelent dolgozatokra [4], a disszertációkra [3] és a gép program leírásokra [2] a következő minta szerint kell hivatkozni:

- [1] Farkas, J., Über die Theorie der einfachen Ungleichungen, *Journal für die reine und angewandte Mathematik* 124 (1902) 1-27.
- [2] Kéri, G., „DUALSIMP” rutin a CDC 3300-as gépekre. (Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutató Intézete, CDC 3300 felhasználói ismertető 2. 1973. május) 19-20.
- [3] Prékopa, A., „Sztochasztikus rendszerek optimalizálási problémáiról”, doktori értekezés. Magyar Tudományos Akadémia, Budapest, 1970.
- [4] Prabhu, N. U., „Recent research on the ruin problem of collective risk theory”, in: *Inventory Control and Water Storage* Ed. A. Prékopa (János Bolyai Mathematical Society and North-Holland Publishing Company, Amsterdam-London, 1973) 221-228.
- [5] Zoutendijk, G., *Methods of Feasible Directions* (Elsevier Publishing Company, Amsterdam and New York, 1960).

A dolgozatok szövegében az irodalmi hivatkozás számait szögletes zárójelben kell megadni, mint például [5] vagy [4, 76-78]. A szerzők a dolgozatukról 50 darab ingyenes különlenyomatot kapnak. A dolgozatok után szerzői díjat az Alkalmazott Matematikai Lapok nem fizet.

Ára: 425 Ft

TARTALOMJEGYZÉK

<i>Bíró M., Mayer J., Rapcsák T. és Vermes M., Matematikai programozási szakértői rendszerekről</i>	217
<i>Mészáros Csaba, Bázisinverz nagyméretű, ritkás lineáris programozás feladatokban</i>	279
<i>Wiandt Tamás, Lineáris tömegpontrendszer csillapított mozgásairól</i>	311
<i>Prok István, Kocka alaptartományú euklideszi kristálycsoportok</i>	321
<i>Molnár Sándor, Szidarovszky Ferenc és Okuguchi Koji, Nemlineáris differenciaegyenletek globális aszimptotikus stabilitásának általános kritériumairól</i>	339
<i>A külföldi szakirodalomból</i>	
<i>Varga, R. S., Matematikai problémák és sejtések numerikus megközelítésben</i>	349

INDEX

<i>Bíró, M., Mayer, J., Rapcsák, T. and Vermes, M., On expert systems of mathematical programming</i>	217
<i>Mészáros, Cs., Basis inverse of large-scale and sparse linear programming problems</i>	279
<i>Wiandt, T., On damped oscillation of linear mechanical systems</i>	311
<i>Prok, I., Euclidean crystallographic groups with cubic fundamental domain</i>	321
<i>Bálint, E. és Deák, I., Parallel computers: optimization software</i>	339
<i>Molnár, S., Szidarovszky, F. és Okuguchi, K., On general criteria for the global asymptotic stability of nonlinear difference equations</i>	357
<i>Csendes, T., An interval method for bounding level sets</i>	367
<i>From the foreign literature</i>	
<i>Varga, R. S., Scientific Computation on Mathematical Problems and Conjectures</i>	389

✓ 317. 471

ALKALMAZOTT MATEMATIKAI LAPOK

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI ÉS FIZIKAI
TUDOMÁNYOK OSZTÁLYÁNAK KÖZLEMÉNYEI

ALAPÍTOTTÁK

KALMÁR LÁSZLÓ, TANDORI KÁROLY, PRÉKOPA ANDRÁS, ARATÓ MÁTYÁS

FŐSZERKESZTŐ

BENCZÚR ANDRÁS

FŐSZERKESZTŐ-HELYETTESEK

DEMETROVICS JÁNOS, FARKAS MIKLÓS

FELELŐS SZERKESZTŐ

SZÁNTAI TAMÁS

A SZERKESZTŐBIZOTTSÁG TAGJAI

ARATÓ MÁTYÁS, CSIRIK JÁNOS, CSISZÁR IMRE, GALÁNTAI AURÉL,
GÉCSEG FERENC, GYIRES BÉLA, GYÖRFFY LÁSZÓ, HARNOS ZSOLT,
HATVANI LÁSZLÓ, HEPPES ANDRÁS, KÁTAI IMRE, KATONA GYULA, KIS OTTÓ,
KLAFSZKY EMIL, KOVÁCS MARGIT, LOVÁSZ LÁSZLÓ, MAROS ISTVÁN,
PRÉKOPA ANDRÁS, RECSKI ANDRÁS, STOYAN GISBERT,
TANDORI KÁROLY, TUSNÁDY GÁBOR, VARGA LÁSZLÓ

XVI. KÖTET

ELTE, BUDAPEST

1992

TARTALOMJEGYZÉK

<i>Bíró M., Mayer J., Rapcsák T. és Vermes M., Matematikai programozási szakértői rendszerekről</i>	217
<i>Deák Jenő, Tóth János és Vizvári Béla, Anyagmegmaradás összetett kémiai mechanizmusokban</i>	73
<i>Domokos Gábor, Síkbeli, konzervatív tenzormezők forgásszimmetriáinak vizsgálata elemi módszerekkel</i>	1
<i>Gregorics Tibor, Megjegyzések a heurisztikus gráftereső algoritmusokhoz</i>	61
<i>Lámer Géza, A szükséges és elégséges összeférhetőségi peremfeltételek meghatározása</i>	99
<i>Lengyel Imre, Megjegyzések a Gilmore–Gomory módszer gyakorlati alkalmazásához</i>	45
<i>Mayer J., Bíró M., Rapcsák T. és Vermes M., Matematikai programozási szakértői rendszerekről</i>	217
<i>Mészáros Csaba, Bázisinverz nagyméretű, ritkás lineáris programozás feladatokban</i>	279
<i>Molnár Sándor, Szidarovszky Ferenc és Okuguchi Koji, Nemlineáris differenciaegyenletek globális aszimptotikus stabilitásának általános kritériumairól</i>	339
<i>Okuguchi Koji, Molnár Sándor és Szidarovszky Ferenc, Nemlineáris differenciaegyenletek globális aszimptotikus stabilitásának általános kritériumairól</i>	339
<i>Prok István, Kocka alaptartományú euklideszi kristálycsoportok</i>	321
<i>Rapcsák T., Bíró M., Mayer J. és Vermes M., Matematikai programozási szakértői rendszerekről</i>	217
<i>Szemők Árpád, Optimális ragadozási stratégia a ragadozó–zsákmány viszonyban</i>	27
<i>Szidarovszky Ferenc, Molnár Sándor és Okuguchi Koji, Nemlineáris differenciaegyenletek globális aszimptotikus stabilitásának általános kritériumairól</i>	339
<i>Tóth János, Deák Jenő és Vizvári Béla, Anyagmegmaradás összetett kémiai mechanizmusokban</i>	73
<i>Tusnády Gábor, A rákkutatás matematikai alapjai</i>	115
<i>Vermes M., Bíró M., Mayer J. és Rapcsák T., Matematikai programozási szakértői rendszerekről</i>	217
<i>Vizvári Béla, Deák Jenő és Tóth János, Anyagmegmaradás összetett kémiai mechanizmusokban</i>	73
<i>Wiandt Tamás, Lineáris tömegpontrendszer csillapított mozgásairól</i>	311
A külföldi szakirodalomból	
<i>Saari, D. G., Káosz és a választások elmélete</i>	131
<i>Varga, R. S., Matematikai problémák és sejtések numerikus megközelítésben</i>	143
<i>Varga, R. S., Matematikai problémák és sejtések numerikus megközelítésben</i>	349
Könyvismertetés	215

INDEX

<i>Bíró, M., Mayer, J., Rapcsák, T. and Vermes, M., On expert systems of mathematical programming</i>	217
<i>Deák, J., Tóth, J. and Vizvári, B., Conservation of matter in complex chemical mechanisms</i>	73
<i>Domokos, G., Investigation of the rotational symmetries of planar, conservative tensor fields by elementary methods</i>	1
<i>Gregorics, T., Remarks to the heuristic graph-search algorithms</i>	61
<i>Lámer, G., Determination of necessary and sufficient compatibility boundary conditions</i>	99
<i>Lengyel, I., Computational notes to the Gilmore-Gomory cutting stock algorithm</i>	45
<i>Mayer, J., Bíró, M., Rapcsák, T. and Vermes, M., On expert systems of mathematical programming</i>	217
<i>Mészáros, Cs., Basis inverse of large-scale and sparse linear programming problems</i>	279
<i>Molnár, S., Szidarovszky, F. és Okuguchi, K., On general criteria for the global asymptotic stability of nonlinear difference equations</i>	339
<i>Okuguchi, K., Molnár, S. és Szidarovszky, F., On general criteria for the global asymptotic stability of nonlinear difference equations</i>	339
<i>Prok, I., Euclidean crystallographic groups with cubic fundamental domain</i>	321
<i>Rapcsák, T., Bíró, M., Mayer, J. and Vermes, M., On expert systems of mathematical programming</i>	217
<i>Szemők, Á., Optimal predator functional response</i>	27
<i>Szidarovszky, F., Molnár, S. és Okuguchi, K., On general criteria for the global asymptotic stability of nonlinear difference equations</i>	339
<i>Tóth, J., Deák, J. and Vizvári, B., Conservation of matter in complex chemical mechanisms</i>	73
<i>Tusnádi, G., Mathematical methods of cancer research</i>	115
<i>Vermes, M., Bíró, M., Mayer, J. and Rapcsák, T., On expert systems of mathematical programming</i>	217
<i>Vizvári, B., Deák, J. és Tóth, J., Conservation of matter in complex chemical mechanisms</i>	73
<i>Wiandt, T., On damped oscillation of linear mechanical systems</i>	311
<i>From the foreign literature</i>	
<i>Saari, D. G., Chaos and the theory of elections</i>	131
<i>Varga, R. S., Scientific Computation on Mathematical Problems and Conjectures</i>	143
<i>Varga, R. S., Scientific Computation on Mathematical Problems and Conjectures</i>	349
<i>Book review</i>	215

ELTE 94115